

(19) 世界知的所有権機関  
国際事務局(43) 国際公開日  
2004 年 11 月 11 日 (11.11.2004)

PCT

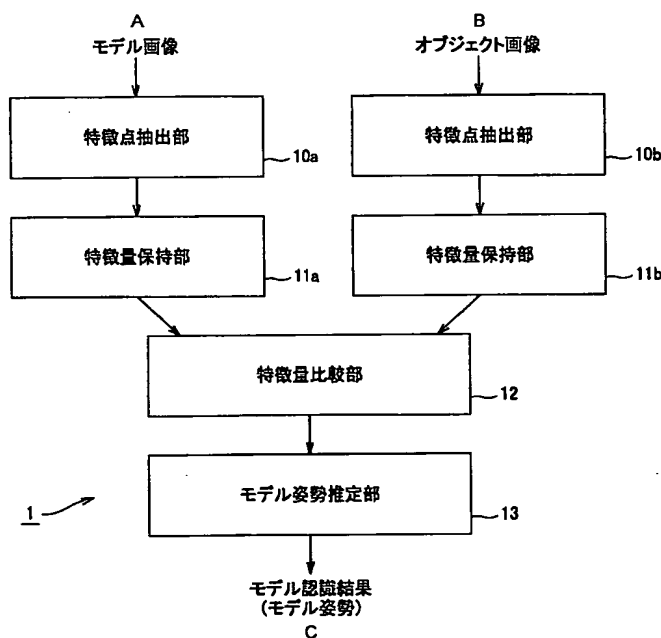
(10) 国際公開番号  
WO 2004/097742 A1

- (51) 国際特許分類<sup>7</sup>: G06T 7/00 (72) 発明者; および  
(21) 国際出願番号: PCT/JP2004/005784 (75) 発明者/出願人 (米国についてのみ): 鈴木 洋貴  
(22) 国際出願日: 2004 年 4 月 22 日 (22.04.2004) (SUZUKI, Hirotsuka) [JP/JP]; 〒1410001 東京都品川区  
(25) 国際出願の言語: 日本語 北品川 6 丁目 7 番 3 5 号 ソニー株式会社内 Tokyo  
(26) 国際公開の言語: 日本語 (JP). 佐部 浩太郎 (SABE, Kohtaro) [JP/JP]; 〒1410001  
(30) 優先権データ: 特願2003-124225 2003 年 4 月 28 日 (28.04.2003) JP 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式  
(71) 出願人 (米国を除く全ての指定国について): ソニー 株式会社 (SONY CORPORATION) [JP/JP]; 〒1410001 東京都品川区北品川 6 丁目 7 番 3 5 号 Tokyo (JP). 株式会社内 Tokyo (JP). 藤田 雅博 (FUJITA, Masahiro) [JP/JP]; 〒1410001 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社内 Tokyo (JP).  
(74) 代理人: 小池 晃, 外 (KOIKE, Akira et al.); 〒1000011 東京都千代田区内幸町一丁目 1 番 7 号 大和生命ビル 1 1 階 Tokyo (JP).

[続葉有]

(54) Title: IMAGE RECOGNITION DEVICE AND METHOD, AND ROBOT DEVICE

(54) 発明の名称: 画像認識装置及び方法、並びにロボット装置



A...MODEL IMAGE  
B...OBJECT IMAGE  
10a...CHARACTERISTIC POINT EXTRACTION SECTION  
10b...CHARACTERISTIC POINT EXTRACTION SECTION  
11a...CHARACTERISTIC AMOUNT HOLDING SECTION  
11b...CHARACTERISTIC AMOUNT HOLDING SECTION  
12...CHARACTERISTIC AMOUNT COMPARISON SECTION  
13...MODEL POSTURE ESTIMATION SECTION  
C...MODEL RECOGNITION RESULT (MODEL POSTURE)

(57) Abstract: In an image recognition device (1), characteristic point extraction sections (10a, 10b) extract a characteristic point from a model image and an object image while characteristic amount holding sections (11a, 11b) extract a characteristic amount for each of the characteristic points and holds them together with the characteristic point position information. A characteristic amount comparison section (12) compares the characteristic amounts by one to another so as to calculate similarity or difference degree and generates a pair of candidate corresponding characteristic points having a high correspondence possibility. A model posture estimation section (13) repeatedly projects to a parameter space the affin conversion parameter decided from three pairs selected at random from the candidate corresponding characteristic point pairs, defines each member of a cluster having the maximum number of members formed on the parameter space as an inlier, calculates an affin conversion parameter by minimum square estimation using the inlier, and outputs the model posture decided by this affin conversion parameter.

(57) 要約: 画像認識装置 (1) において、特徴点抽出部 (10a), (10b) は、モデル画像及びオブジェクト画像から特徴点を抽出し、特徴量保持部 (11a), (11b) は、各特徴点について特徴量を抽出し、特徴点の位置情報と共に保持する。特徴量比較部 (12) は、特徴量同士を比較して類似度又は相違度を計算し、対応する可能性の高い候補対応特徴点ペアを生成する。モデル姿勢推定部 (13) は、候補対応特徴点ペア群からランダムに選択した 3 ペアから決定される

アフィン変換パラメータをパラメータ空間に投射する操作を繰り返し、パラメータ空間上で形成された最多メンバ数を持つクラスターの各メンバをインライヤとし、インライヤを用いた最小自乗推定によりアフィン変換パラメータを求め、

[続葉有]



(81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:

— 国際調査報告書

(84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL,

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

## 明細書

## 画像認識装置及び方法、並びにロボット装置

## 技術分野

本発明は、複数のオブジェクトを含むオブジェクト画像から検出対象となるモデルを抽出する画像認識装置及びその方法、並びにそのような画像認識機能を搭載したロボット装置に関する。

本出願は、日本国において2003年4月28日に出願された日本特許出願番号2003-124225を基礎として優先権を主張するものであり、この出願は参照することにより、本出願に援用される。

## 背景技術

現在、実用化されている対象物認識技術の多くは、残差逐次検定法や相互相関係数によるテンプレートマッチング手法を用いている。しかしながら、このテンプレートマッチング手法は、検出対象物体が入力画像中に変形なく出現すると仮定できる特殊な場合には有効であるが、視点や照明状態の一定でない一般的な画像からの対象物認識環境においては有効でない。

また、検出対象物体の形状特徴と画像分割手法によって切り出された入力画像中の各領域の形状特徴とをマッチングする形状マッチング手法も提案されているが、上述のような一般的な対象物認識環境においては領域分割の結果が安定せず、入力画像中の物体の良質な形状記述が難しくなる。特に、検出対象物体が他の物体に部分的に隠されている場合などは認識が非常に困難になる。

さらに、以上のような入力画像或いはその部分領域の全体的な特徴を用いたマッチング手法に対し、入力画像から特徴的な点（特徴点）やエッジを抽出し、それらが構成する線分集合やエッジ集合の空間的位置関係を線図形やグラフとして表現し、線図形間或いはグラフ間の構造類似性に基づいてマッチングを行う方法

も提案されている。しかしながら、この手法はある特化した対象物に対しては上手く作用するが、画像の変形により安定した特徴点間構造が抽出されない場合があり、特に上述したような部分的に隠される場合などにおける認識は困難となる。

一方で、画像から特徴点を抽出し、特徴点とその局所近傍の画像情報から得られる特徴量を用いたマッチング手法も存在する。例えば、C. シュミット及びR. ムーアは、文献「C. Schmid and R. Mohr, "Local grayvalue invariants for image retrieval", IEEE PAMI, Vol. 19, No. 5, pp. 530-534, 1997」(以下、「文献1」という。)において、Harrisコーナー検出器を用いて検出されたコーナーを特徴点とし、その特徴点付近の回転不変特徴量を用いたマッチング手法を提案している。このような特徴点の部分的画像変形に対して不変な局所特徴量を用いるマッチング手法では、上述の手法と比較して画像の変形に対しても検出対象が部分的に隠されるような場合にも安定した検出が可能となる。しかしながら、この文献1で用いられている特徴量は、画像の拡大縮小変換に対して不変性を持たないものであるため、拡大縮小変換のある場合には認識が困難となる。

これに対して、D. ロウは、文献「D. Lowe, "Object recognition from local scale-invariant features", Proc. of the International Conference on Computer Vision, Vol. 2, pp. 1150-1157, September 20-25, 1999, Corfu, Greece」(以下、「文献2」という。)において、画像の拡大縮小変換に対しても不変であるような特徴点及び特徴量を用いたマッチング手法を提案している。このD. ロウの提案した画像認識装置について、図1を用いて説明する。

図1に示す画像認識装置400において、特徴点抽出部401a、401bは、特徴点抽出対象画像(モデル画像又はオブジェクト画像)から得られた画像の多重解像度表現(スケールスペース表現。文献「Lindeberg T., "Scale-space: A framework for handling image structures at multiple scales.", Journal of Applied Statistics, Vol. 21, No. 2, pp. 224-270, 1994」参照)の各解像度画像に対してDOG(Difference of Gaussian)フィルタを適用し、DOGフィルタ出力画像の局所点(局所極大点及び局所極小点)のうち、所定の範囲内の解像度変化によって位置の変化がないような点を特徴点として検出する。ここで、解像度の階層数は予め設定しておく。

続いて特徴量保持部 402 a, 402 b は、特徴点抽出部 401 a, 401 b にて抽出された各特徴点について特徴量を抽出し、保持する。この際、特徴点抽出部 401 a, 401 b は、特徴点近傍領域のカノニカル・オリエンテーション (canonical orientation; 支配的方向) と、オリエンテーション・プレーン (orientation plane) とを用いる。ここで、カノニカル・オリエンテーションは、ガウス重み付き勾配強度を累積した方向ヒストグラムのピーク値を与える方向であり、特徴量保持部 402 a, 402 b は、このカノニカル・オリエンテーションを特徴量として保持する。また、特徴量保持部 402 a, 402 b は、特徴点近傍領域の勾配強度情報をカノニカル・オリエンテーションで正規化、すなわちカノニカル・オリエンテーションを 0 deg として方向補正し、近傍領域内各点の勾配強度情報を位置情報と共に勾配方向で分類する。例えば、45 deg 毎の全部で 8 つのオリエンテーション・プレーンに近傍領域内各点の勾配強度情報を分類する場合、近傍領域の局所座標系上の点 (x, y) における方向 93 deg、強度 m の勾配情報は、90 deg のラベルを持つと共に近傍領域と同一の局所座標系を持つオリエンテーション・プレーンの位置 (x, y) に強度 m の情報としてマッピングされる。その後、各オリエンテーション・プレーンは、解像度のスケールに応じたぼかし及びリサンプリングが施される。特徴量保持部 402 a, 402 b は、このようにして求めた (解像度数) × (オリエンテーション・プレーン数) × (各オリエンテーション・プレーンのサイズ) 次元の特徴量ベクトルを保持する。

続いて、特徴量比較部 403 は、各オブジェクト特徴点の特徴量と最も特徴量の類似するモデル特徴点を k-d tree 法 (検索効率のよい特徴空間上の Nearest Neighbor 探索法) を用いて検索し、得られた候補対応特徴点ペアを候補対応特徴点ペア群として保持する。

一方で、モデル姿勢粗推定部 404 は、一般化ハフ変換により、モデル特徴点とオブジェクト特徴点との空間的位置関係からモデルのオブジェクト画像上での姿勢 (回転角度, 拡大縮小率, 平行移動量の画像変換パラメータ) を推定する。この際、一般化ハフ変換のパラメータ参照テーブル (R テーブル) のインデックスとして、上述した各特徴点のカノニカル・オリエンテーションが用いられると

予想される。このモデル姿勢粗推定部 404 の出力は、画像変換パラメータ空間上への投票結果であり、最多投票数を獲得したパラメータがモデル姿勢の粗い推定を与える。

そこで、候補対応特徴点ペア選択部 405 は、そのパラメータに投票を行ったオブジェクト特徴点をメンバに持つ候補対応特徴点ペアのみを選択することにより、候補対応特徴点ペア群の絞り込みを行う。

最後に、モデル姿勢推定部 406 は、「検出されるモデルはオブジェクト画像上へのアフィン変換による画像変形が施されている」という拘束条件の下、対応特徴点ペア群の空間的配置から最小自乗推定によりアフィン変換パラメータを推定する。そして、モデル姿勢推定部 406 は、候補対応特徴点ペア群の各モデル特徴点をこのアフィン変換パラメータによりオブジェクト画像上に変換し、対応するオブジェクト特徴点との位置のずれ（空間的距離）を求め、このずれが著しいペアを排除し、候補対応特徴点ペア群を更新する。この際、モデル姿勢推定部 406 は、候補対応特徴点ペア群が 2 組以下の場合には「モデル検出不可」と出力して終了し、そうでなければ所定の終了条件が満たされるまでこの操作を繰り返す。終了条件を満たした際のアフィン変換パラメータで決定されるモデル姿勢をモデル認識結果として出力する。

しかしながら、この文献 2 に記載された D. ロウの手法には幾つかの問題点がある。

まず第 1 に、特徴点におけるカノニカル・オリエンテーションの抽出に問題がある。上述したように、カノニカル・オリエンテーションは、特徴点近傍領域の局所勾配情報から求めた、ガウス重み付き勾配強度を累積した方向ヒストグラムのピーク値を与える方向により求まる。ここで、文献 2 の手法では、物体のコーナーの若干内側に特徴点が検出される傾向があるが、このような特徴点の近傍の方向ヒストグラムには、2 つのピークがそれぞれエッジに直交する方向に出現するため、複数の競合するカノニカル・オリエンテーションが検出されうるが、後段の特徴量比較部 403 及びモデル姿勢粗推定部 404 ではこのような場合を想定しておらず、対処できない。また、ガウス重み関数のパラメータに依存して方向ヒストグラムの形状が変化し、カノニカル・オリエンテーションの安定した抽

出ができないという問題点もある。その一方で、このカノニカル・オリエンテーションは、後段の特徴量比較部 403 及びモデル姿勢粗推定部 404 で用いられるため、不適切なカノニカル・オリエンテーションの抽出は特徴量マッチングの結果に重大な影響を与えてしまう。

第 2 に、オリエンテーション・プレーンによる特徴量比較において、局所領域内各点の濃度勾配強度情報による特徴量マッチングを行っているが、一般に勾配強度は明度の変化に対して不変な特徴量ではないため、モデル画像とオブジェクト画像との間に明度差がある場合には、安定なマッチングが保証されないという問題がある。

第 3 に、各オブジェクト特徴点に対して、特徴空間上での距離が最短ではないが十分小さい、すなわち特徴量が十分類似するモデル特徴点が複数存在し、その中に真の特徴点ペア（インライヤ）が含まれている場合が想定されるが、特徴量比較部 403 においては、各オブジェクト特徴点は特徴空間上で最短距離を与えるモデル特徴点のみとペアが組まれるため、先のようなインライヤが候補対応ペアとして考慮されていない点も問題として挙げられる。

第 4 に、モデル姿勢推定部 406 におけるアフィン変換パラメータ推定の際に問題が起こり得る。ここで、候補対応特徴点ペア選択部 405 で絞り込みを受けた対応特徴点ペア群の中には偽の特徴点ペア（アウトライヤ）が含まれているが、候補対応特徴点ペア群中のアウトライヤが多い場合や真のアフィン変換パラメータから極端に逸脱したアウトライヤが存在する場合には、アフィン変換パラメータ推定はアウトライヤに影響を受け、場合によっては繰り返し操作によってインライヤが次第に排除されてアウトライヤが残り、誤ったモデル姿勢を出力してしまうという問題も起こり得る。

## 発明の開示

本発明は、このような従来の実情に鑑みて提案されたものであり、複数の物体を含み、物体同士が部分的に重なり合っているような画像からも対象物を検出可能であり、さらに視点の変化（平行移動、拡大縮小、回転、ストレッチを含む画

像変化)、明度変化、ノイズによる画像情報の変形がある場合であっても安定に対象物を検出可能である画像認識装置及びその方法、並びにそのような画像認識機能を搭載したロボット装置を提供することを目的とする。

上述した目的を達成するために、本発明に係る画像認識装置及びその方法は、複数のオブジェクトを含むオブジェクト画像と検出対象となるモデルを含むモデル画像とを比較し、上記オブジェクト画像から上記モデルを抽出するものであり、上記オブジェクト画像及び上記モデル画像の各々から特徴点を抽出する特徴点抽出手段(工程)と、上記オブジェクト画像及び上記モデル画像の各々について、少なくとも上記特徴点の近傍領域における濃度勾配情報から得られる濃度勾配方向ヒストグラムを特徴量として抽出して保持する特徴量保持手段(工程)と、上記オブジェクト画像の各特徴点と上記モデル画像の各特徴点とを比較し、類似する特徴量を有する候補対応特徴点ペアを生成する特徴量比較手段(工程)と、上記候補対応特徴点ペアを用いて上記オブジェクト画像上の上記モデルの有無を検出し、上記モデルが存在する場合に該モデルの位置及び姿勢を推定するモデル姿勢推定手段(工程)とを有し、上記特徴量比較手段(工程)は、比較対象となる特徴点同士の上記濃度勾配方向ヒストグラムの一方を濃度勾配方向に巡回的にシフトさせながら該濃度勾配方向ヒストグラム間の距離を求め、最小の距離を該濃度勾配方向ヒストグラム間の距離として、上記候補対応特徴点ペアを生成するものである。

このような画像認識装置及びその方法では、特徴点の近傍領域における濃度勾配情報から得られる濃度勾配方向ヒストグラムを特徴量として特徴量マッチングを行う際に、比較対象となる特徴点同士の上記濃度勾配方向ヒストグラムの一方を濃度勾配方向に巡回的にシフトさせながら該濃度勾配方向ヒストグラム間の距離を求め、最小の距離を該濃度勾配方向ヒストグラム間の距離として、距離の類似する特徴点同士で候補対応特徴点ペアを生成する。

また、上述した目的を達成するために、本発明に係る画像認識装置及びその方法は、複数のオブジェクトを含むオブジェクト画像と検出対象となるモデルを含むモデル画像とを比較し、上記オブジェクト画像から上記モデルを抽出するものであり、上記オブジェクト画像及び上記モデル画像の各々から特徴点を抽出する



特徴点抽出手段（工程）と、上記オブジェクト画像及び上記モデル画像の各々について、上記特徴点の近傍領域における特徴量を抽出して保持する特徴量保持手段（工程）と、上記オブジェクト画像の各特徴点と上記モデル画像の各特徴量とを比較し、類似する特徴量を有する候補対応特徴点ペアを生成する特徴量比較手段（工程）と、上記候補対応特徴点ペアを用いて上記オブジェクト画像上の上記モデルの有無を検出し、上記モデルが存在する場合に該モデルの位置及び姿勢を推定するモデル姿勢推定手段（工程）とを有し、上記モデル姿勢推定手段（工程）は、ランダムに選択した３組の候補対応特徴点ペアから決定されるアフィン変換パラメータをパラメータ空間に投射する操作を繰り返し、パラメータ空間上で形成されたクラスタのうち最多メンバ数を持つクラスタに属するアフィン変換パラメータに基づいて上記モデルの位置及び姿勢を決定するアフィン変換パラメータを求めるものである。

このような画像認識装置及びその方法では、特徴量が類似するとして生成された候補対応特徴点ペアを用いて上記オブジェクト画像上のモデルの有無を検出し、モデルが存在する場合に該モデルの位置及び姿勢を推定する際に、ランダムに選択した３組の候補対応特徴点ペアから決定されるアフィン変換パラメータをパラメータ空間に投射する操作を繰り返し、パラメータ空間上で形成されたクラスタのうち最多メンバ数を持つクラスタに属するアフィン変換パラメータに基づいて上記モデルの位置及び姿勢を決定するアフィン変換パラメータを求める。

また、本発明に係るロボット装置は、上述したような画像認識機能を搭載したものである。

本発明のさらに他の目的、本発明によって得られる具体的な利点は、以下に説明される実施例の説明から一層明らかにされるであろう。

#### 図面の簡単な説明

図１は、従来の画像認識装置の概略構成を示す図である。

図２は、本実施の形態における画像認識装置の概略構成を示す図である。

図３は、同画像認識装置の特徴点抽出部における画像の多重解像度ピラミッド

構造の構築方法を説明する図である。

図 4 は、第 1 レベルまでの解像度変化によって位置の変化がないような特徴点の検出処理を説明するフローチャートである。

図 5 は、第 3 レベルまでの解像度変化によって位置の変化がないような特徴点を検出する場合の例を示す図である。

図 6 A 及び図 6 B は、同画像認識装置の特徴量保持部における処理を説明する図であり、図 6 A は、特徴点から半径 3.5 ピクセルの範囲内を近傍構造とした場合における特徴点近傍の濃度勾配情報の例を示し、図 6 B は、図 6 A の濃度勾配情報から得られる勾配方向ヒストグラムの例を示す。

図 7 は、同画像認識装置の特徴量比較部における処理の詳細を説明するフローチャートである。

図 8 は、濃度勾配ベクトル  $U_m$ ,  $U_o$  間の類似度の計算手法を説明する図である。

図 9 は、同画像認識装置のモデル姿勢推定部における処理の詳細を説明するフローチャートである。

図 10 は、候補対応特徴点ペア選択部を有する画像認識装置の概略構成を示す図である。

図 11 A 乃至図 11 C は、同画像認識装置の候補対応特徴点ペア選択部における第 1 の手法を説明する図であり、図 11 A は、候補対応特徴点ペア群の例を示し、図 11 B は、各候補対応特徴点ペアに保持されている推定回転角度を示し、図 11 C は、推定回転角度ヒストグラムを示す。

図 12 は、本実施の形態におけるロボット装置の外観構成を示す斜視図である。

図 13 は、同ロボット装置の自由度構成モデルを模式的に示す図である。

図 14 は、同ロボット装置のシステム構成を示す図である。

発明を実施するための最良の形態

以下、本発明を適用した具体的な実施の形態について、図面を参照しながら詳細に説明する。この実施の形態は、本発明を、複数のオブジェクトを含む入力画像であるオブジェクト画像と検出対象となるモデルを含むモデル画像とを比較し、オブジェクト画像からモデルを抽出する画像認識装置に適用したものである。

先ず、本実施の形態における画像認識装置の概略構成を図2に示す。図2に示す画像認識装置1において、特徴点抽出部10a, 10bは、モデル画像及びオブジェクト画像からそれぞれモデル特徴点及びオブジェクト特徴点を抽出し、特徴量保持部11a, 11bは、抽出された各特徴点について後述する特徴量を抽出し、特徴点の位置情報と共に保持する。そして特徴量比較部12は、各モデル特徴点の特徴量と各オブジェクト特徴点の特徴量とを比較して類似度又は相違度を計算し、この類似尺度を用いて特徴量が類似する、すなわち対応する可能性の高いモデル特徴点及びオブジェクト特徴点のペア（候補対応特徴点ペア）を生成する。

モデル姿勢推定部13は、生成された候補対応特徴点ペア群を用いてオブジェクト画像上のモデルの有無を検出し、「モデル有」の判定結果の場合には、「検出されるモデルはオブジェクト画像上へのアフィン変換による画像変形が施されている」という拘束条件の下、候補対応特徴点ペア群からランダムに選択した3ペアにより決定されるアフィン変換パラメータをパラメータ空間に投射する操作を繰り返し、パラメータ空間上で形成されたクラスタのうち最多メンバ数を持つクラスタの各メンバを真の特徴点ペア（インライヤ）とし、このインライヤを用いた最小自乗推定によりアフィン変換パラメータを求める。そして、モデル姿勢推定部13は、このアフィン変換パラメータで決定されるモデル姿勢をモデル認識結果として出力する。

以下、上記画像認識装置1の各ブロックについて詳細に説明する。なお、以下では、画像の水平方向をX軸、垂直方向をY軸として説明する。

特徴点抽出部10a, 10bは、特徴点抽出対象画像（モデル画像又はオブジェクト画像）に対し、先ず平滑化フィルタリング、例えば以下の式（1）で示す2次元ガウス関数による畳み込み積（ガウスフィルタリング）と、双2次線形補間リサンプリングによる画像縮小とを繰り返し交互に適用することで、画像の多重解像度ピラミッド構造を構築する。ここで、リサンプリングファクタとしては、式（1）のガウスフィルタで用いた $\sigma$ を用いる。

$$g(x, y) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad \dots (1)$$

すなわち、図3に示すように、例えば入力画像  $I$  に  $\sigma = \sqrt{2}$  であるガウスフィルタ  $g(x, y)$  を施すことにより第1レベル（最高解像度）の画像  $I_1$  を生成し、さらにガウスフィルタを施すことにより画像  $g * I_1$  を生成する。そして、この画像  $g * I_1$  をリサンプリングしてガウスフィルタを施すことにより第2レベルの画像  $I_2$ 、 $g * I_2$  を生成し、同様にして画像  $g * I_2$  から第3レベルの画像  $I_3$ 、 $g * I_3$  を生成する。

続いて、特徴点抽出部 10a, 10b は、各レベル（各解像度）の画像に対して、DOG（Difference of Gaussian）フィルタを適用する。このDOGフィルタは、画像の輪郭強調のために用いられる2次微分フィルタの一種であり、人間の視覚系で網膜からの情報が外側膝状体で中継されるまでに行われている処理の近似モデルとして、LOG（Laplacian of Gaussian）フィルタと共によく用いられるものである。DOGフィルタの出力は、2つのガウスフィルタ出力画像の差分を取ることで容易に得られる。すなわち、図3に示すように、第1レベルの画像について画像  $DI_1 (= I_1 - g * I_1)$  が得られ、第2レベル、第3レベルの画像について画像  $DI_2 (= I_2 - g * I_2)$ 、画像  $DI_3 (= I_3 - g * I_3)$  が得られる。

そして、特徴点抽出部 10a, 10b は、各レベルにおけるDOGフィルタ出力画像  $DI_1$ 、 $DI_2$ 、 $DI_3 \dots$  の局所点（局所極大点及び局所極小点）のうち、所定の範囲内の解像度変化によって位置の変化がないような点を特徴点として検出する。これにより、画像の拡大縮小操作に対してロバストな特徴点間のマッチングが実現できる。

ここで、多重解像度ピラミッド構造の第  $L$  レベル、すなわち  $\sigma$  の  $(L-1)$  乗ファクタまでの解像度変化によって位置の変化がないような特徴点を検出する際の処理について図4のフローチャートを用いて説明する。

先ずステップS1において、第1レベル（最高解像度）におけるD o Gフィルタ出力画像D I<sub>1</sub>の局所点（局所極大点及び局所極小点）を検出する。なお、局所近傍としては、例えば3×3直接近傍を用いることができる。

次にステップS2において、検出された各局所点について、その1つ上位レベル（解像度の1つ低い層）における対応点を解像度減少に伴う画像縮小を考慮して求め、この対応点が局所点か否かを判別する。局所点である場合（Yes）にはステップS3に進み、局所点でない場合（No）には探索を終了する。

続いてステップS3では、第Lレベルまで探索が成功したか否かを判別する。第Lレベルまで探索を行っていない場合（No）にはステップS2に戻ってさらに上位レベルの探索を行い、第Lレベルまで探索が成功した場合（Yes）には特徴点であるとして、ステップS4においてその位置情報を保持する。

例えば第3レベルまでの解像度変化によって位置の変化がないような特徴点を検出する場合、図5に示すように、第1レベルの画像D I<sub>1</sub>において検出された局所点F P<sub>1</sub>、F P<sub>2</sub>のうち、F P<sub>1</sub>は第3レベルの画像D I<sub>3</sub>まで対応点が存在するため特徴点とされ、F P<sub>2</sub>は第2レベルまでしか対応点が存在しないため特徴点でないとされる。

なお、この特徴点抽出部10a、10bでは、D o Gフィルタの代わりにL o Gフィルタを用いるようにしても構わない。また、文献「Harris C. and Stephens M., "A combined corner and edge detector.", in Proc. Alvey Vision Conf., pp. 147-151, 1988」で物体のコーナー検出に用いられているcorner-ness関数の出力値をD o Gフィルタ出力の代わりに用いるようにしても構わない。

次に、特徴量保持部11a、11b（図2）は、特徴点抽出部10a、10bにて抽出された各特徴点について特徴量を抽出し、保持する。特徴量としては、多重解像度ピラミッド構造の各レベルの画像（I<sub>l</sub>, l = 1, ..., L）の画像情報から導出される特徴点の近傍領域各点の濃度勾配情報（勾配強度及び勾配方向）を用いる。点（x、y）における勾配強度M<sub>x,y</sub>及び勾配方向R<sub>x,y</sub>は以下の式（2）、（3）により与えられる。

$$M_{xy} = \sqrt{(I_{x+1,j} - I_{xy})^2 + (I_{xy+1} - I_{xy})^2} \quad \dots (2)$$

$$R_{xy} = \tan^{-1}(I_{xy+1} - I_{xy}, I_{x+1,y} - I_{xy}) \quad \dots (3)$$

ここで、特徴量を算出する特徴点近傍領域としては、回転変化に対して構造の変わらないような、特徴点に関して対称なものを選択することが好ましい。これにより、回転変化に対するロバスト性を実現することができる。例えば、(i) 特徴点から半径  $r$  ピクセルの範囲内を特徴点近傍領域とする手法や、(ii) 特徴点を中心に幅  $\sigma$  の特徴点に関して対称な 2 次元ガウス重みを濃度勾配に掛ける手法を用いることができる。

特徴点  $FP$  から半径 3.5 ピクセルの範囲内を近傍領域とした場合における特徴点近傍領域の濃度勾配情報の例を図 6 A に示す。ここで、図 6 A において矢印の長さは勾配強度を表し、矢印の方向は勾配方向を表す。

また、特徴量保持部 11 a, 11 b は、特徴点近傍の勾配方向に関するヒストグラム（方向ヒストグラム）も特徴量として保持する。図 6 A の濃度勾配情報から得られる勾配方向ヒストグラムの例を図 6 B に示す。ここで、図 6 B における階級幅  $\Delta \theta$  は  $10 \text{ deg}$  であり、階級数  $N$  は 36 ( $= 360 \text{ deg} / 10 \text{ deg}$ ) である。

続いて、特徴量比較部 12（図 2）は、各モデル特徴点の特徴量と各オブジェクト特徴点の特徴量とを比較し、類似する特徴量を有するモデル特徴点及びオブジェクト特徴点のペア（候補対応特徴点ペア）を生成する。

この特徴量比較部 12 における処理の詳細について、図 7 のフローチャートを用いて説明する。まず、ステップ S 10 において、特徴量比較部 12 は、各モデル特徴点の方向ヒストグラムと各オブジェクト特徴点の方向ヒストグラムとを比較してヒストグラム間の距離（相違度）を計算すると共に、モデルーオブジェクト間の推定回転角度を求める。

ここで、階級幅  $\Delta \theta$  と階級数  $N$  とが同一である 2 つの方向ヒストグラム  $H_1 =$

$\{h_1(n), n=1, \dots, N\}$  及び  $H_2 = \{h_2(n), n=1, \dots, N\}$  を想定し、 $h_1(n)$ 、 $h_2(n)$  が階級  $n$  における度数を示すものとする、ヒストグラム  $H_1$  とヒストグラム  $H_2$  との間の距離  $d(H_1, H_2)$  は、例えば以下の式 (4) で与えられる。ここで、式 (4) における  $r$  としては、 $r=1, 2, \infty$  が一般的に用いられる。

$$d(H_1, H_2) = \left( \sum_i \|h_1(i) - h_2(i)\|^r \right)^{1/r} \quad \dots (4)$$

この式 (4) を用いて各モデル特徴点及び各オブジェクト特徴点の方向ヒストグラム間の相違度を計算するが、(i) モデル-オブジェクト間のスケール比がマッチング段階で未知であるため、モデル特徴点の各レベルとオブジェクト特徴点の各レベル間で方向ヒストグラム間のマッチングを行う必要がある。また、(ii) 方向ヒストグラム間のマッチングに関しては、モデル-オブジェクト間の回転変換量を考慮する必要がある。

モデル特徴点  $m$  のレベル  $L^v$  における方向ヒストグラム  $H_m^{L^v} = \{h_m^{L^v}(n), n=1, \dots, N\}$  と、オブジェクト特徴点  $o$  のレベル  $l^v$  における方向ヒストグラム  $H_o^{l^v} = \{h_o^{l^v}(n), n=1, \dots, N\}$  との相違度を求める場合を考える。方向ヒストグラムは回転変換に対して巡回的に変化するため、 $H_o^{l^v}$  を巡回的に階級を 1 つずつシフトさせながら式 (4) の計算を行い、その最小値を  $H_m^{L^v}$  と  $H_o^{l^v}$  との間の相違度とする。このとき、相違度の最小値を与えたときのシフト量 (ずらした階級数) からオブジェクト特徴点の回転角度を推定することができる。なお、この手法は方向ヒストグラム交差法として知られている。

$H_o^{l^v}$  を  $k$  階級分シフトさせた方向ヒストグラムを  $H_o^{l^v(k)}$  とすると、方向ヒストグラム交差法による方向ヒストグラム間相違度 dissimilarity ( $H_m^{L^v}, H_o^{l^v(k)}$ ) は、以下の式 (5) で与えられる。

$$\text{dissimilarity}(H_m^{LV}, H_o^{lv}) = \min_{k=0}^{N-1} (d(H_m^{LV}, H_o^{lv(k)})) \quad \dots (5)$$

また、最小の  $d(H_m^{LV}, H_o^{lv(k)})$  を与える  $k$  を  $k'$  とすると、オブジェクト特徴点  $o$  の近傍領域における推定回転角度  $\theta(m, LV, o, lv)$  は以下の式 (6) で与えられる。

$$\theta(m, LV, o, lv) = k' \Delta\theta \quad \dots (6)$$

上述の (i) を考慮すると、モデル特徴点  $m$  とオブジェクト特徴点  $o$  との方向ヒストグラム間相違度  $\text{dissimilarity}(H_m, H_o)$  は、以下の式 (7) のようになる。

$$\text{dissimilarity}(H_m, H_o) = \min_{LV, lv} (\text{dissimilarity}(H_m^{LV}, H_o^{lv})) \quad \dots (7)$$

特徴量比較部 12 は、モデル特徴点  $m$  とオブジェクト特徴点  $o$  との各ペア ( $m, n$ ) に対し、最小の方向ヒストグラム間相違度  $\text{dissimilarity}(H_m, H_o)$  を与えるレベル  $LV, lv$  (以下、それぞれ  $LV_m^*$ ,  $lv_o^*$  と表記する。) と、その推定回転角度  $\theta(m, LV_m^*, o, lv_o^*)$  とを、方向ヒストグラム間相違度  $\text{dissimilarity}(H_m, H_o)$  と共に保持する。

次にステップ S11 (図7) において、特徴量比較部 12 は、各モデル特徴点  $m$  に対し方向ヒストグラム間相違度の小さい順に  $K$  個のオブジェクト特徴点  $o_{m1}$ ,



...,  $o_{mk}$ を選択し、候補対応特徴点ペアを組ませる。すなわち、各モデル特徴点  $m$  に対して  $K$  個の候補対応特徴点ペア  $(m, o_{m1}), \dots, (m, o_{mk}), \dots, (m, o_{mk})$  が組まれる。また、各候補対応特徴点ペア  $(m, o_{mk})$  には対応するレベル  $L V_m^*$ ,  $l v_{o_{mk}}^*$  と推定回転角度  $\theta(m, L V_m^*, o, l v_{o_{mk}}^*)$  との情報が保持される。

以上のようにして、全モデル特徴点に対し候補対応特徴点ペアを組ませることで得られたペア群が、候補対応特徴点ペア群となる。

このように、特徴量比較部 12 では、ヒストグラム度数に勾配強度を累積しておらず、単に勾配方向のみに注目しているため、明度変化に対してロバストな特徴量マッチングが可能になる。また、前述した文献 2 の手法では、カノニカル・オリエンテーションのように抽出が不安定な特徴量を元にマッチングを行っているのに対して、本実施の形態では方向ヒストグラムの形状を考慮したより安定なマッチングを行うことができる。また、二次的に安定した特徴量（推定回転角度）を得ることができる。

なお、上述のステップ S 11 では、各モデル特徴点  $m$  に対して  $K$  個の候補対応特徴点ペアを選択するものとして説明したが、これに限定されるものではなく、方向ヒストグラム間相違度が閾値を下回るペア全てを選択するようにしても構わない。

ここで、上述の操作で生成された候補対応特徴点ペア群は、方向ヒストグラムが類似するものの濃度勾配の空間的特徴が異なっているような対応点ペアも含んでいる。そこで、続くステップ S 12（図 7）では、濃度勾配ベクトル間類似度によるペアの選択、候補対応特徴点ペア群の更新を行う。

具体的には、モデル特徴点  $m$  の近傍のレベル  $L V_m^*$  における濃度勾配ベクトルを  $U_m$  とし、このモデル特徴点  $m$  と対応点ペアを組むオブジェクト特徴点  $o$  の近傍のレベル  $l v_{o_{mk}}^*$  における濃度勾配ベクトルを  $U_o$  としたとき、 $U_m$  と  $U_o$  との類似度が閾値を下回るペアを排除し、候補対応特徴点ペア群を更新する。

濃度勾配ベクトル  $U_m$ ,  $U_o$  間の類似度の計算手法を図 8 に示す。まず、 $U_m$  を空間的に 4 領域  $R_i$  ( $i = 1, \dots, 4$ ) に分割し、各領域の平均濃度勾配ベクトル  $\nabla_i$  ( $i = 1, \dots, 4$ ) を求める。この  $\nabla_i$  をまとめた 8 次元ベクトル  $\nabla$  で  $U_m$  が

表現される。一方、回転変換を考慮した濃度勾配情報のマッチングを行うために、 $U_0$ の勾配方向を先に求めた推定回転角度  $\theta$  ( $m, L V_m^*, 0, l v_{omk}^*$ ) で補正し、 $U_0^*$ を得る。この際、中間位置の値は、双2次線形補間により求める。先と同様に  $U_0^*$ を4領域  $R_i$  ( $i = 1, \dots, 4$ ) に分割し、各領域の平均濃度勾配ベクトル  $\mathbf{V}_i$  ( $i = 1, \dots, 4$ ) を求める。この  $\mathbf{V}_i$ をまとめた8次元ベクトル  $\mathbf{V}$  で  $U_0$  が表現される。このとき、 $U_m$  と  $U_0$  との類似度  $\text{similarity}(U_m, U_0) \in [0, 1]$  は、平均濃度勾配ベクトル  $\mathbf{V}$  と  $\mathbf{W}$  との類似度として解釈され、例えばコサイン相関値を用いて以下の式 (8) により求められる。ここで、式 (8) において  $(\mathbf{V} \cdot \mathbf{W})$  は  $\mathbf{V}$  と  $\mathbf{W}$  との内積を表す。

$$\text{similarity}(U_m, U_0) = \frac{1}{2} \left( \frac{(\mathbf{V} \cdot \mathbf{W})}{\|\mathbf{V}\| \|\mathbf{W}\|} + 1 \right) \quad \dots (8)$$

特徴量比較部 12 は、各候補対応特徴点ペアに対して上記式 (8) で得られる平均濃度勾配ベクトル間の類似度を求め、類似度が閾値  $\delta$  を下回るペアを候補対応特徴点ペア群から排除し、候補対応特徴点ペア群を更新する。

このように、特徴量比較部 12 では、部分領域の平均濃度勾配ベクトルを用いて特徴量の比較を行っているため、特徴点位置や推定回転角度の微妙なずれや、明度の変化による濃度勾配情報の変化に対してロバストなマッチングを実現することができ、計算量の削減にもなる。

以上の操作により、特徴点近傍の局所的な濃度勾配情報が類似する (モデル特徴点-オジェクト特徴点) ペア群を抽出することができるが、巨視的に見ると、このように得られたペア群は、対応特徴点間の空間的位置関係がモデルのオブジェクト画像上での姿勢 (モデル姿勢) と矛盾するような「偽の特徴点ペア (アウトライヤ)」を含んでしまっている。

ここで、候補対応特徴点ペアが3組以上ある場合には、最小自乗推定により近似アフィン変換パラメータの推定が可能であり、推定モデル姿勢と空間的位置関

係の矛盾する対応ペアを排除し、残ったペアで再びモデル姿勢推定を行うという操作を繰り返すことで、モデル姿勢を認識することができる。

しかしながら、候補対応特徴点ペア群中のアウトライヤが多い場合や真のアフィン変換パラメータから極端に逸脱したアウトライヤが存在する場合には、最小自乗推定による推定結果は一般的に満足のいくものではないことが知られている (Hartley R., Zisserman A., "Multiple View Geometry in Computer Vision.", Chapter 3, pp. 69-116, Cambridge University Press, 2000) ため、本実施の形態におけるモデル姿勢推定部 13 (図 2) では、アフィン変換拘束の下、候補対応特徴点ペア群の空間的位置関係から「真の特徴点ペア (インライヤ)」を抽出し、抽出されたインライヤを用いてモデル姿勢 (平行移動量, 回転, 拡大縮小, ストレッチを決定するアフィン変換パラメータ) を推定する。

以下、このモデル姿勢推定部 13 における処理について説明する。上述したように、候補対応特徴点ペアが 3 組以上なければアフィン変換パラメータを決定できないため、候補対応特徴点ペアが 2 組以下の場合、モデル姿勢推定部 13 は、オブジェクト画像中にモデルが存在しない、又はモデル姿勢検出に失敗したとして、「認識不可」と出力し処理を終了する。一方、候補対応特徴点ペアが 3 組以上ある場合、モデル姿勢推定部 13 は、モデル姿勢を検出可能であるとし、アフィン変換パラメータの推定を行う。なお、モデル姿勢推定部 13 は、モデル画像及びオブジェクト画像の例えば第 1 レベル (最高解像度) における特徴点の空間的位置を元にモデル姿勢を推定するものとする。

ここで、モデル特徴点  $[x \ y]^T$  のオブジェクト特徴点  $[u \ v]^T$  へのアフィン変換は、以下の式 (9) で与えられる。

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \quad \dots (9)$$

この式 (9) において、 $a_i$  ( $i = 1, \dots, 4$ ) は回転, 拡大縮小, ストレッチ

を決定するパラメータを表し、 $[b_1 \ b_2]^T$ は平行移動パラメータを表す。決定すべきアフィン変換パラメータは $a_1, \dots, a_4$ 及び $b_1, b_2$ の6つであるため、候補対応特徴点ペアが3組あればアフィン変換パラメータを決定することができる。

3組の候補対応特徴点ペアで構成されるペア群  $P$  を  $([x_1 \ y_1]^T, [u_1 \ v_1]^T)$ ,  $([x_2 \ y_2]^T, [u_2 \ v_2]^T)$ ,  $([x_3 \ y_3]^T, [u_3 \ v_3]^T)$  とすると、ペア群  $P$  とアフィン変換パラメータとの関係は、以下の式 (10) に示す線形システムで表現することができる。

$$\begin{bmatrix} x_1 & y_1 & 0 & 0 & 1 & 0 \\ 0 & 0 & x_1 & y_1 & 0 & 1 \\ x_2 & y_2 & 0 & 0 & 1 & 0 \\ 0 & 0 & x_2 & y_2 & 0 & 1 \\ x_3 & y_3 & 0 & 0 & 1 & 0 \\ 0 & 0 & x_3 & y_3 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} u_1 \\ v_1 \\ u_2 \\ v_2 \\ u_3 \\ v_3 \end{bmatrix} \quad \dots (10)$$

この式 (10) を  $A \mathbf{x} = \mathbf{b}$  のように書き直すと、アフィン変換パラメータ  $\mathbf{x}$  の最小自乗解は以下の式 (11) で与えられる。

$$\mathbf{x} = A^{-1} \mathbf{b} \quad \dots (11)$$

さて、候補対応特徴点ペア群からアウトライヤが1つ以上混入するように繰り返しランダムにペア群  $P$  を選択した場合、そのアフィン変換パラメータはパラメータ空間上に散らばって投射される。一方、インライヤのみから構成されるペア群  $P$  を繰り返しランダムに選択した場合、そのアフィン変換パラメータは、何れもモデル姿勢の真のアフィン変換パラメータに極めて類似した、すなわちパ

ラメータ空間上で距離の近いものとなる。したがって、候補対応特徴点ペア群からランダムにペア群  $P$  を選択し、そのアフィン変換パラメータをパラメータ空間上に投射していく操作を繰り返すと、インライヤはパラメータ空間上で密度の高い（メンバ数の多い）クラスタを形成し、アウトライヤは散らばって出現することになる。このことから、パラメータ空間上でクラスタリングを行い、最多メンバ数を持つクラスタの要素がインライヤとなる。

モデル姿勢推定部 13 における処理の詳細を図 9 のフローチャートを用いて説明する。なお、このモデル姿勢推定部 13 におけるクラスタリング手法としては、NN (Nearest Neighbor) 法を用いるものとする。この際、上述した  $b_1$ ,  $b_2$  は、認識対象画像により様々な値を取り得るため、 $\mathbf{x}$  空間でもクラスタリングにおいてクラスタリング閾値の選択が認識対象に依存してしまう。そこで、モデル姿勢推定部 13 では、「真のパラメータと  $a_1, \dots, a_4$  は類似するが  $b_1, b_2$  が異なるようなアフィン変換パラメータを与えるペア群  $P$  は殆ど存在しない」という仮定の下、パラメータ  $a_1, \dots, a_4$  (以下、 $\mathbf{a}$  と表記する。) で張られるパラメータ空間上のみでクラスタリングを行う。なお、上記仮定が成り立たない状況が生じたとしても、 $\mathbf{a}$  空間とは独立に  $b_1, b_2$  で張られるパラメータ空間でクラスタリングを行いその結果を考慮することで、容易に問題を回避することができる。

先ず図 9 のステップ S 20 において初期化を行う。具体的には、繰り返し数のカウント値  $cnt$  を  $cnt = 1$  とし、候補対応特徴点ペア群からランダムにペア群  $P_1$  を選択し、アフィン変換パラメータ  $\mathbf{a}_1$  を求める。また、クラスタ数  $N$  を  $N = 1$  とし、アフィン変換パラメータ空間  $\mathbf{a}$  上で  $\mathbf{a}_1$  を中心とするクラスタ  $C_1$  を作る。このクラスタ  $C_1$  のセントロイド  $\mathbf{c}_1$  を  $\mathbf{c}_1 = \mathbf{a}_1$  とし、メンバ数  $nc_1$  を  $nc_1 = 1$  とし、カウント値  $cnt$  を  $cnt = 2$  に更新する。

次にステップ S 21 において、モデル姿勢推定部 13 は、候補対応特徴点ペア群からランダムにペア群  $P_{cnt}$  を選択し、アフィン変換パラメータ  $\mathbf{a}_{cnt}$  を求める。

続いてステップ S 22 において、モデル姿勢推定部 13 は、NN 法を用いてアフィン変換パラメータ空間のクラスタリングを行う。具体的には、先ず以下の式

(12)に従ってアフィン変換パラメータ  $a_{cnt}$  と各クラスタ  $C_i$  のセントロイド  $c_i$  ( $i=1, \dots, N$ ) との距離  $d(a_{cnt}, c_i)$  のうち、最小の距離  $d_{min}$  を求める。

$$d_{min} = \min_{1 \leq i \leq N} \{d(a_{cnt}, c_i)\} \quad \dots (12)$$

そして、所定の閾値  $\tau$  (例えば  $\tau=0.1$ ) に対して  $d_{min} < \tau$  であれば  $d_{min}$  を与えるクラスタ  $C_i$  に  $a_{cnt}$  を属させ、 $a_{cnt}$  を含めた全メンバでクラスタ  $C_i$  のセントロイド  $c_i$  を更新する。また、クラスタ  $C_i$  のメンバ数  $nc_i$  を  $nc_i = nc_i + 1$  とする。一方、 $d_{min} \geq \tau$  であれば、アフィン変換パラメータ空間  $a$  上で  $a_{cnt}$  をセントロイド  $c_{N+1}$  とする新しいクラスタ  $C_{N+1}$  を作り、メンバ数  $nc_{N+1}$  を  $nc_{N+1} = 1$  とし、クラスタ数  $N$  を  $N = N + 1$  とする。

続いてステップ S23 では、繰り返し終了条件を満たすか否かが判別される。ここで、繰り返し終了条件としては、例えば最多メンバ数が所定の閾値 (例えば 15) を超え、且つ最多メンバ数と 2 番目に多いメンバ数との差が所定の閾値 (例えば 3) を超える場合、或いは繰り返し数カウンタのカウント値  $cnt$  が所定の閾値 (例えば 5000 回) を超える場合に終了するように設定することができる。ステップ S23 において、繰り返し終了条件を満たさない場合 (No) には、ステップ S24 で繰り返し数のカウント値  $cnt$  を  $cnt = cnt + 1$  とした後、ステップ S21 に戻る。一方、繰り返し終了条件を満たす場合 (Yes) にはステップ S25 に進む。

最後にステップ S25 において、モデル姿勢推定部 13 は、以上で得られたインライヤを用いて、最小自乗法によりモデル姿勢を決定するアフィン変換パラメータを推定する。

ここで、インライヤを  $([x_{in1} \ y_{in1}]^T, [u_{in1} \ v_{in1}]^T)$ ,  $([x_{in2} \ y_{in2}]^T, [u_{in2} \ v_{in2}]^T)$ , ... とすると、インライヤとアフィン変換パラメータとの関係は、以下の式 (13) に示す線形システムで表現することができる。

$$\begin{bmatrix} x_{IN1} & y_{IN1} & 0 & 0 & 1 & 0 \\ 0 & 0 & x_{IN1} & y_{IN1} & 0 & 1 \\ x_{IN2} & y_{IN2} & 0 & 0 & 1 & 0 \\ 0 & 0 & x_{IN2} & y_{IN2} & 0 & 1 \\ \dots & & & & & \\ \dots & & & & & \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} u_{IN1} \\ v_{IN1} \\ u_{IN2} \\ v_{IN2} \\ \vdots \\ \vdots \end{bmatrix} \quad \dots (13)$$

この式(13)を $\mathbf{A}_{IN} \mathbf{x}_{IN} = \mathbf{b}_{IN}$ のように書き直すと、アフィン変換パラメータ $\mathbf{x}_{IN}$ の最小自乗解は以下の式(14)で与えられる。

$$\mathbf{x}_{IN} = (\mathbf{A}_{IN}^T \mathbf{A}_{IN})^{-1} \mathbf{A}_{IN}^T \mathbf{b}_{IN} \quad \dots (14)$$

そしてステップS25では、このアフィン変換パラメータ $\mathbf{x}_{IN}$ で決定されるモデル姿勢をモデル認識結果として出力する。

なお、以上の説明では、閾値 $\tau$ が定数値であるものとしたが、ステップS21乃至ステップS24の繰り返し処理を行う際に、始めは比較的大きな閾値 $\tau$ を用いて大雑把なインライヤ抽出を行い、繰り返し回数が増える毎に次第に小さい閾値 $\tau$ を用いる、いわゆる「焼きなまし法」のような手法を適用してもよい。これにより、精度よくインライヤを抽出することができる。

また、以上の説明では、候補対応特徴点ペア群からランダムにペア群 $P$ を選択し、そのアフィン変換パラメータをパラメータ空間上に投射していく操作を繰り返し、パラメータ空間上で最多メンバ数を持つクラスタの要素をインライヤとして、最小自乗法によりモデル姿勢を決定するアフィン変換パラメータを推定したが、これに限定されるものではなく、例えば最多メンバ数を持つクラスタのセン

トロイドを、モデル姿勢を決定するアフィン変換パラメータとしても構わない。

ところで、特徴量比較部 12 で生成された候補対応特徴点ペア群中のアウトライヤの比率が大きくなるほどモデル姿勢推定部 13 におけるインライヤの選択確率が低下し、モデル姿勢を推定する際に多くの繰り返し回数が必要となるため、計算時間が増大してしまう。したがって、この検出可能判定部 13 に入力される候補対応特徴点ペア群からできる限りアウトライヤを排除しておくことが望ましい。そこで、本実施の形態における画像認識装置 1 では、図 10 に示すように、特徴量比較部 12 とモデル姿勢推定部 13 との間に以下に説明するような候補対応特徴点ペア選択部 14 を追加することができる。

この候補対応特徴点ペア選択部 14 は、第 1 の手法として、推定回転角度ヒストグラムを作成することにより候補対応特徴点ペアを選択する。具体的に、図 11 A に示すようなモデル  $m_d$  を含むモデル画像及びオブジェクト  $ob_1$ 、 $ob_2$  を含むオブジェクト画像を想定して説明する。特徴量比較部 12 では、モデル特徴点  $m$  とオブジェクト特徴点  $o$  との間で、図 11 A に示すような候補対応特徴点ペア群  $P_1, \dots, P_6$  が生成される。このうち、 $P_1, P_2, P_5, P_6$  はインライヤであり、 $P_3, P_4$  はアウトライヤであるものとする。

ここで、特徴量比較部 12 で生成された各候補対応特徴点ペアには、モデルのオブジェクト画像上での推定回転角度情報が保持されているが、図 11 B に示すように、インライヤの推定回転角度が何れも近い値（例えば  $40 \text{ deg}$ ）を有するのに対し、アウトライヤの推定回転角度は様々な値（例えば  $110 \text{ deg}$ 、 $260 \text{ deg}$ ）を有する。したがって、図 11 C に示すような推定回転角度ヒストグラムを作成すると、そのピークを与える推定回転角度を有するペアは、インライヤ（及び推定回転角度がインライヤと一致する極少数のアウトライヤ）となる。

そこで、候補対応特徴点ペア選択部 14 は、特徴量比較部 12 で生成された候補対応特徴点ペア群のうち、推定回転角度ヒストグラムのピークを与える推定回転角度を有するペアを選択し、モデル姿勢推定部 13 に供給する。これにより、安定且つ精密にモデル姿勢のアフィン変換パラメータを推定することが可能となる。ただし、モデルに対するストレッチ変換が著しい場合、画像中の各点における回転角度は一定ではなくなるため、この第 1 の手法は、著しいストレッチ変換



を想定していない場合にのみ有効である。

また、候補対応特徴点ペア選択部 1 4 は、第 2 の手法として、一般化ハフ変換を用いてモデル姿勢の粗推定を行う。具体的には、特徴量比較部 1 2 で生成された候補対応特徴点ペア群について、回転、拡大縮小率、平行移動 (x、y 方向) の 4 つの画像変換パラメータを特徴空間 (投票空間) とした一般化ハフ変換を行う。最も投票の多かった画像変換パラメータ (最多投票パラメータ) によりモデルのオブジェクト画像上での粗い推定モデル姿勢が決定されるが、一方で最多投票パラメータに投票した候補対応特徴点ペア群が、この粗い推定モデル姿勢をサポートするインライヤ (及び極少数のアウトライヤ) となっている。

そこで、候補対応特徴点ペア選択部 1 4 は、最多投票パラメータに投票した候補対応特徴点ペア群をモデル姿勢推定部 1 3 に供給する。これにより、安定且つ精密にモデル姿勢のアフィン変換パラメータを推定することが可能となる。

なお、候補対応特徴点ペア選択部 1 4 は、上述した第 1、第 2 の手法を併用するようにしても構わない。

以上説明したように、本実施の形態における画像認識装置 1 によれば、複数の物体を含み、物体同士が部分的に重なり合っているようなオブジェクト画像からもモデルを検出可能であり、さらに視点の変化 (平行移動、拡大縮小、回転、ストretchを含む画像変化)、明度変化、ノイズによる画像情報の変形に対してもロバストである。

ここで、この画像認識装置 1 は、例えば図 1 2 に示すようなロボット装置に搭載することができる。この図 1 2 に示す 2 足歩行タイプのロボット装置 3 0 は、住環境その他の日常生活上の様々な場面における人的活動を支援する実用ロボットであり、内部状態 (怒り、悲しみ、喜び、楽しみ等) に応じて行動できるほか、人間が行う基本的な動作を表出できるエンターテインメントロボットである。

図 1 2 に示すように、ロボット装置 3 0 は、体幹部ユニット 3 1 の所定の位置に頭部ユニット 3 2 が連結されると共に、左右 2 つの腕部ユニット 3 3 R/L と、左右 2 つの脚部ユニット 3 4 R/L が連結されて構成されている (ただし、R 及び L の各々は、右及び左の各々を示す接尾辞である。以下において同じ。 ) 。

このロボット装置 3 0 が具備する関節自由度構成を図 1 3 に模式的に示す。頭

部ユニット102を支持する首関節は、首関節ヨー軸101と、首関節ピッチ軸102と、首関節ロール軸103という3自由度を有している。

また、上肢を構成する各々の腕部ユニット33R/Lは、肩関節ピッチ軸107と、肩関節ロール軸108と、上腕ヨー軸109と、肘関節ピッチ軸110と、前腕ヨー軸111と、手首関節ピッチ軸112と、手首関節ロール軸113と、手部114とで構成される。手部114は、実際には、複数本の指を含む多関節・多自由度構造体である。ただし、手部114の動作は、ロボット装置1の姿勢制御や歩行制御に対する寄与や影響が少ないので、本明細書ではゼロ自由度と仮定する。したがって、各腕部は7自由度を有するとする。

また、体幹部ユニット2は、体幹ピッチ軸104と、体幹ロール軸105と、体幹ヨー軸106という3自由度を有する。

また、下肢を構成する各々の脚部ユニット34R/Lは、股関節ヨー軸115と、股関節ピッチ軸116と、股関節ロール軸117と、膝関節ピッチ軸118と、足首関節ピッチ軸119と、足首関節ロール軸120と、足部121とで構成される。本明細書中では、股関節ピッチ軸116と股関節ロール軸117の交点は、ロボット装置30の股関節位置を定義する。人体の足部121は、実際には多関節・多自由度の足底を含んだ構造体であるが、ロボット装置30の足底は、ゼロ自由度とする。したがって、各脚部は、6自由度で構成される。

以上を総括すれば、ロボット装置30全体としては、合計で $3 + 7 \times 2 + 3 + 6 \times 2 = 32$ 自由度を有することになる。ただし、エンターテインメント向けのロボット装置30が必ずしも32自由度に限定されるわけではない。設計・制作上の制約条件や要求仕様等に応じて、自由度すなわち関節数を適宜増減することができるというまでもない。

上述したようなロボット装置30が持つ各自由度は、実際にはアクチュエータを用いて実装される。外観上で余分な膨らみを排してヒトの自然体形状に近似させること、2足歩行という不安定構造体に対して姿勢制御を行うことなどの要請から、アクチュエータは小型且つ軽量であることが好ましい。また、アクチュエータは、ギア直結型で且つサーボ制御系をワンチップ化してモータユニット内に搭載したタイプの小型ACサーボ・アクチュエータで構成することがより好まし

い。

図14には、ロボット装置30の制御システム構成を模式的に示している。図14に示すように、制御システムは、ユーザ入力などに動的に反応して情緒判断や感情表現を司る思考制御モジュール200と、アクチュエータ350の駆動などロボット装置1の全身協調運動を制御する運動制御モジュール300とで構成される。

思考制御モジュール200は、情緒判断や感情表現に関する演算処理を実行するCPU (Central Processing Unit) 211や、RAM (Random Access Memory) 212、ROM (Read Only Memory) 213、及び、外部記憶装置 (ハード・ディスク・ドライブなど) 214で構成される、モジュール内で自己完結した処理を行うことができる、独立駆動型の情報処理装置である。

この思考制御モジュール200は、画像入力装置251から入力される画像データや音声入力装置252から入力される音声データなど、外界からの刺激などに従って、ロボット装置30の現在の感情や意思を決定する。ここで、画像入力装置251は、例えばCCD (Charge Coupled Device) カメラを複数備えており、また、音声入力装置252は、例えばマイクロホンを複数備えている。

また、思考制御モジュール200は、意思決定に基づいた動作又は行動シーケンス、すなわち四肢の運動を実行するように、運動制御モジュール300に対して指令を発行する。

一方の運動制御モジュール300は、ロボット装置30の全身協調運動を制御するCPU 311や、RAM 312、ROM 313、及び外部記憶装置 (ハード・ディスク・ドライブなど) 314で構成される、モジュール内で自己完結した処理を行うことができる、独立駆動型の情報処理装置である。外部記憶装置314には、例えば、オフラインで算出された歩行パターンや目標とするZMP軌道、その他の行動計画を蓄積することができる。ここで、ZMPとは、歩行中の床反力によるモーメントがゼロとなる床面上の点のことであり、また、ZMP軌道とは、例えばロボット装置30の歩行動作期間中にZMPが動く軌跡を意味する。なお、ZMPの概念並びにZMPを歩行ロボットの安定度判別規範に適用する点については、Miomir Vukobratovic 著 “LEGGED LOCOMOTION ROBOTS” (加藤一郎

外著『歩行ロボットと人工の足』（日刊工業新聞社））に記載されている。

運動制御モジュール 300 には、図 13 に示したロボット装置 30 の全身に分散するそれぞれの関節自由度を実現するアクチュエータ 350、体幹部ユニット 2 の姿勢や傾斜を計測する姿勢センサ 351、左右の足底の離床又は着床を検出する接地確認センサ 352、353、バッテリーなどの電源を管理する電源制御装置 354 などの各種の装置が、バス・インターフェース（I/F）301 経由で接続されている。ここで、姿勢センサ 351 は、例えば加速度センサとジャイロ・センサの組み合わせによって構成され、接地確認センサ 352、353 は、近接センサ又はマイクロ・スイッチなどで構成される。

思考制御モジュール 200 と運動制御モジュール 300 は、共通のプラットフォーム上で構築され、両者間はバス・インターフェース 201、301 を介して相互接続されている。

運動制御モジュール 300 では、思考制御モジュール 200 から指示された行動を体現すべく、各アクチュエータ 350 による全身協調運動を制御する。すなわち、CPU 311 は、思考制御モジュール 200 から指示された行動に応じた動作パターンを外部記憶装置 314 から取り出し、又は、内部的に動作パターンを生成する。そして、CPU 311 は、指定された動作パターンに従って、足部運動、ZMP 軌道、体幹運動、上肢運動、腰部水平位置及び高さなどを設定するとともに、これらの設定内容に従った動作を指示する指令値を各アクチュエータ 350 に転送する。

また、CPU 311 は、姿勢センサ 351 の出力信号によりロボット装置 30 の体幹部ユニット 31 の姿勢や傾きを検出するとともに、各接地確認センサ 352、353 の出力信号により各脚部ユニット 5R/L が遊脚又は立脚の何れの状態であるかを検出することによって、ロボット装置 30 の全身協調運動を適応的に制御することができる。

また、CPU 311 は、ZMP 位置が常に ZMP 安定領域の中心に向かうように、ロボット装置 30 の姿勢や動作を制御する。

さらに、運動制御モジュール 300 は、思考制御モジュール 200 において決定された意思通りの行動がどの程度発現されたか、すなわち処理の状況を、思考

制御モジュール 200 に返すようになっている。

このようにしてロボット装置 30 は、制御プログラムに基づいて自己及び周囲の状況を判断し、自律的に行動することができる。

このロボット装置 30 において、上述した画像認識機能をインプリメントしたプログラム（データを含む）は、例えば思考制御モジュール 200 の ROM 213 に格納される。この場合、思考制御モジュール 200 の CPU 211 によって画像認識プログラムが実行される。

このようなロボット装置 30 に上述した画像認識機能を組み込むことにより、ロボット装置 30 は、画像入力装置 251 を介して入力される画像データから予め記憶しているモデルを精度よく抽出することができる。特に、ロボット装置 30 が自律的に歩行している際に画像入力装置 251 の CCD カメラで撮像した周囲の画像から所望のモデルを検出するような場合、他の障害物によってそのモデルが部分的に隠れていることが多々あり、視点や明度も一定ではないが、上述した画像認識手法によれば、このような場合においてもモデルを精度よく抽出することができる。

なお、本発明は、図面を参照して説明した上述の実施例に限定されるものではなく、添付の請求の範囲及びその主旨を逸脱することなく、様々な変更、置換又はその同等のものを行うことができることは当業者にとって明らかである。

#### 産業上の利用可能性

上述した本発明に係る画像認識装置によれば、ヒストグラム度数に勾配強度を累積せず、単に勾配方向のみに注目して候補対応特徴点ペアを生成しているため、明度変化に対してロバストな特徴量マッチングが可能になる。また、方向ヒストグラムの形状を考慮したより安定なマッチングを行うことができると共に、二次的に安定した特徴量（推定回転角度）を得ることができる。

また、本発明に係る画像認識装置によれば、特徴量が類似するとして生成された候補対応特徴点ペアを用いてオブジェクト画像上のモデルの有無を検出し、モデルが存在する場合に該モデルの位置及び姿勢を推定する際に、最小自乗推定を

用いてモデルの位置及び姿勢を決定するアフィン変換パラメータを求めるのではなく、アフィン変換パラメータを投射したパラメータ空間上で最多メンバ数を持つクラスタに属するアフィン変換パラメータに基づいてアフィン変換パラメータを求めることにより、候補対応特徴点ペアに偽の対応点ペア含まれている場合であっても、安定してモデルの位置及び姿勢を推定することができる。

したがって、このような画像認識装置をロボット装置に搭載することで、ロボット装置は、入力される画像データから予め記憶しているモデルを精度よく抽出することができるようになる。

## 請求の範囲

1. 複数のオブジェクトを含むオブジェクト画像と検出対象となるモデルを含むモデル画像とを比較し、上記オブジェクト画像から上記モデルを抽出する画像認識装置であって、

上記オブジェクト画像及び上記モデル画像の各々から特徴点を抽出する特徴点抽出手段と、

上記オブジェクト画像及び上記モデル画像の各々について、少なくとも上記特徴点の近傍領域における濃度勾配情報から得られる濃度勾配方向ヒストグラムを特徴量として抽出して保持する特徴量保持手段と、

上記オブジェクト画像の各特徴点と上記モデル画像の各特徴点とを比較し、類似する特徴量を有する候補対応特徴点ペアを生成する特徴量比較手段と、

上記候補対応特徴点ペアを用いて上記オブジェクト画像上の上記モデルの有無を検出し、上記モデルが存在する場合に該モデルの位置及び姿勢を推定するモデル姿勢推定手段とを備え、

上記特徴量比較手段は、比較対象となる特徴点同士の上記濃度勾配方向ヒストグラム的一方を濃度勾配方向に巡回的にシフトさせながら該濃度勾配方向ヒストグラム間の距離を求め、最小の距離を該濃度勾配方向ヒストグラム間の距離として、上記候補対応特徴点ペアを生成する

ことを特徴とする画像認識装置。

2. 上記特徴量保持手段は、さらに上記近傍領域を複数の部分領域に分割した各部分領域毎の平均濃度勾配ベクトルを上記特徴量として抽出して保持し、

上記特徴量比較手段は、上記比較対象となる特徴点同士の上記濃度勾配方向ヒストグラム間の距離と、上記部分領域毎の平均濃度勾配ベクトルを上記近傍領域でまとめた特徴ベクトル間の類似度とに基づいて、上記候補対応特徴点ペアを生成する

ことを特徴とする請求の範囲第1項記載の画像認識装置。

3. 上記特徴量比較手段は、上記比較対象となる特徴点同士の上記濃度勾配方向ヒストグラム間の距離に基づき暫定的な候補対応特徴点ペアを生成し、上記特徴

ベクトル間の類似度に基づき該暫定的な候補対応特徴点ペアから上記候補対応特徴点ペアを選択することを特徴とする請求の範囲第2項記載の画像認識装置。

4. 上記特徴量比較手段は、上記最小の距離を与えるシフト量に相当する回転角度で上記近傍領域における濃度勾配ベクトルの濃度勾配方向を補正し、補正後の近傍領域における上記特徴ベクトル間の類似度に基づき上記暫定的な候補対応特徴点ペアから上記候補対応特徴点ペアを選択することを特徴とする請求の範囲第3項記載の画像認識装置。

5. 上記モデル姿勢推定手段は、ランダムに選択した3組の候補対応特徴点ペアから決定されるアフィン変換パラメータをパラメータ空間に投射する操作を繰り返し、パラメータ空間上で形成されたクラスタのうち最多メンバ数を持つクラスタに属するアフィン変換パラメータに基づいて上記モデルの位置及び姿勢を決定するアフィン変換パラメータを求めることを特徴とする請求の範囲第1項記載の画像認識装置。

6. 上記モデル姿勢推定手段は、上記最多メンバ数を持つクラスタのセントロイドを上記モデルの位置及び姿勢を決定するアフィン変換パラメータとすることを特徴とする請求の範囲第5項記載の画像認識装置。

7. 上記モデル姿勢推定手段は、上記最多メンバ数を持つクラスタに属するアフィン変換パラメータを与えた候補対応特徴点ペアを真の候補対応特徴点ペアとし、該真の候補対応特徴点ペアを用いて最小自乗推定により上記モデルの位置及び姿勢を決定するアフィン変換パラメータを求めることを特徴とする請求の範囲第5項記載の画像認識装置。

8. 上記最小の距離を与えるシフト量に相当する回転角度に関する回転角度ヒストグラムを作成し、上記特徴量比較手段によって生成された上記候補対応特徴点ペアから上記回転角度ヒストグラムのピークにおける回転角度を与えた候補対応特徴点ペアを選択する候補対応特徴点ペア選択手段をさらに備え、

上記モデル姿勢推定手段は、上記候補対応特徴点ペア選択手段によって選択された候補対応特徴点ペアを用いて上記オブジェクト画像上の上記モデルの有無を検出し、上記モデルが存在する場合に該モデルの位置及び姿勢を推定する

ことを特徴とする請求の範囲第1項記載の画像認識装置。



9. 上記特徴量比較手段によって生成された候補対応特徴点ペアに対して、回転角度、拡大縮小率並びに水平及び垂直方向の平行移動量をパラメータ空間とした一般化ハフ変換を行い、上記特徴量比較手段によって生成された候補対応特徴点ペアから最多投票パラメータに投票した候補対応特徴点ペアを選択する候補対応特徴点ペア選択手段をさらに備え、

上記モデル姿勢推定手段は、上記候補対応特徴点ペア選択手段によって選択された候補対応特徴点ペアを用いて上記オブジェクト画像上の上記モデルの有無を検出し、上記モデルが存在する場合に該モデルの位置及び姿勢を推定する

ことを特徴とする請求の範囲第1項記載の画像認識装置。

10. 上記特徴点抽出手段は、上記オブジェクト画像又は上記モデル画像に対して平滑化フィルリングと縮小リサンプリングとを繰り返し適用して得られる多重解像度ピラミッド構造において、各解像度における2次微分系フィルタ出力画像の局所極大点又は局所極小点のうち、所定の範囲内の解像度変化によって位置の変化がない点を上記特徴点として抽出することを特徴とする請求の範囲第1項記載の画像認識装置。

11. 複数のオブジェクトを含むオブジェクト画像と検出対象となるモデルを含むモデル画像とを比較し、上記オブジェクト画像から上記モデルを抽出する画像認識装置であって、

上記オブジェクト画像及び上記モデル画像の各々から特徴点を抽出する特徴点抽出手段と、

上記オブジェクト画像及び上記モデル画像の各々について、上記特徴点の近傍領域における特徴量を抽出して保持する特徴量保持手段と、

上記オブジェクト画像の各特徴点と上記モデル画像の各特徴点とを比較し、類似する特徴量を有する候補対応特徴点ペアを生成する特徴量比較手段と、

上記候補対応特徴点ペアを用いて上記オブジェクト画像上の上記モデルの有無を検出し、上記モデルが存在する場合に該モデルの位置及び姿勢を推定するモデル姿勢推定手段とを備え、

上記モデル姿勢推定手段は、ランダムに選択した3組の候補対応特徴点ペアから決定されるアフィン変換パラメータをパラメータ空間に投射する操作を繰り返

し、パラメータ空間上で形成されたクラスタのうち最多メンバ数を持つクラスタに属するアフィン変換パラメータに基づいて上記モデルの位置及び姿勢を決定するアフィン変換パラメータを求める

ことを特徴とする画像認識装置。

12. 上記モデル姿勢推定手段は、上記最多メンバ数を持つクラスタのセントロイドを上記モデルの位置及び姿勢を決定するアフィン変換パラメータとすることを特徴とする請求の範囲第11項記載の画像認識装置。

13. 上記モデル姿勢推定手段は、上記最多メンバ数を持つクラスタに属するアフィン変換パラメータを与えた候補対応特徴点ペアを真の候補対応特徴点ペアとし、該真の候補対応特徴点ペアを用いて最小自乗推定により上記モデルの位置及び姿勢を決定するアフィン変換パラメータを求めることを特徴とする請求の範囲第11項記載の画像認識装置。

14. 上記特徴量比較手段によって生成された候補対応特徴点ペアに対して、回転角度、拡大縮小率並びに水平及び垂直方向の平行移動量をパラメータ空間とした一般化ハフ変換を行い、上記特徴量比較手段によって生成された候補対応特徴点ペアから最多投票パラメータに投票した候補対応特徴点ペアを選択する候補対応特徴点ペア選択手段をさらに備え、

上記モデル姿勢推定手段は、上記候補対応特徴点ペア選択手段によって選択された候補対応特徴点ペアを用いて上記オブジェクト画像上の上記モデルの有無を検出し、上記モデルが存在する場合に該モデルの位置及び姿勢を推定する

ことを特徴とする請求の範囲第11項記載の画像認識装置。

15. 上記特徴点抽出手段は、上記オブジェクト画像又は上記モデル画像に対して平滑化フィルリングと縮小リサンプリングとを繰り返し適用して得られる多重解像度ピラミッド構造において、各解像度における2次微分系フィルタ出力画像の局所極大点又は局所極小点のうち、所定の範囲内の解像度変化によって位置の変化がない点を上記特徴点として抽出することを特徴とする請求の範囲第11項記載の画像認識装置。

16. 複数のオブジェクトを含むオブジェクト画像と検出対象となるモデルを含むモデル画像とを比較し、上記オブジェクト画像から上記モデルを抽出する画像

認識方法であって、

上記オブジェクト画像及び上記モデル画像の各々から特徴点を抽出する特徴点抽出工程と、

上記オブジェクト画像及び上記モデル画像の各々について、少なくとも上記特徴点の近傍領域における濃度勾配情報から得られる濃度勾配方向ヒストグラムを特徴量として抽出して保持する特徴量保持工程と、

上記オブジェクト画像の各特徴点と上記モデル画像の各特徴点とを比較し、類似する特徴量を有する候補対応特徴点ペアを生成する特徴量比較工程と、

上記候補対応特徴点ペアを用いて上記オブジェクト画像上の上記モデルの有無を検出し、上記モデルが存在する場合に該モデルの位置及び姿勢を推定するモデル姿勢推定工程とを有し、

上記特徴量比較工程では、比較対象となる特徴点同士の上記濃度勾配方向ヒストグラム的一方を濃度勾配方向に巡回的にシフトさせながら該濃度勾配方向ヒストグラム間の距離を求め、最小の距離を該濃度勾配方向ヒストグラム間の距離として、上記候補対応特徴点ペアを生成する

ことを特徴とする画像認識方法。

17. 複数のオブジェクトを含むオブジェクト画像と検出対象となるモデルを含むモデル画像とを比較し、上記オブジェクト画像から上記モデルを抽出する画像認識方法であって、

上記オブジェクト画像及び上記モデル画像の各々から特徴点を抽出する特徴点抽出工程と、

上記オブジェクト画像及び上記モデル画像の各々について、上記特徴点の近傍領域における特徴量を抽出して保持する特徴量保持工程と、

上記オブジェクト画像の各特徴点と上記モデル画像の各特徴点とを比較し、類似する特徴量を有する候補対応特徴点ペアを生成する特徴量比較工程と、

上記候補対応特徴点ペアを用いて上記オブジェクト画像上の上記モデルの有無を検出し、上記モデルが存在する場合に該モデルの位置及び姿勢を推定するモデル姿勢推定工程とを有し、

上記モデル姿勢推定工程では、ランダムに選択した3組の候補対応特徴点ペア

から決定されるアフィン変換パラメータをパラメータ空間に投射する操作を繰り返し、パラメータ空間上で形成されたクラスタのうち最多メンバ数を持つクラスタに属するアフィン変換パラメータに基づいて上記モデルの位置及び姿勢を決定するアフィン変換パラメータを求める

ことを特徴とする画像認識方法。

18. 入力画像と検出対象となるモデルを含むモデル画像とを比較し、上記入力画像から上記モデルを抽出することが可能な自律型のロボット装置であって、

外部環境を撮像し、上記入力画像とする画像入力手段と、

上記入力画像及び上記モデル画像の各々から特徴点を抽出する特徴点抽出手段と、

上記入力画像及び上記モデル画像の各々について、少なくとも上記特徴点の近傍領域における濃度勾配情報から得られる濃度勾配方向ヒストグラムを特徴量として抽出して保持する特徴量保持手段と、

上記入力画像の各特徴点と上記モデル画像の各特徴点とを比較し、類似する特徴量を有する候補対応特徴点ペアを生成する特徴量比較手段と、

上記候補対応特徴点ペアを用いて上記入力画像上の上記モデルの有無を検出し、上記モデルが存在する場合に該モデルの位置及び姿勢を推定するモデル姿勢推定手段とを備え、

上記特徴量比較手段は、比較対象となる特徴点同士の上記濃度勾配方向ヒストグラムの一方を濃度勾配方向に巡回的にシフトさせながら該濃度勾配方向ヒストグラム間の距離を求め、最小の距離を該濃度勾配方向ヒストグラム間の距離として、上記候補対応特徴点ペアを生成する

ことを特徴とするロボット装置。

19. 入力画像と検出対象となるモデルを含むモデル画像とを比較し、上記入力画像から上記モデルを抽出することが可能な自律型のロボット装置であって、

外部環境を撮像し、上記入力画像とする画像入力手段と、

上記入力画像及び上記モデル画像の各々から特徴点を抽出する特徴点抽出手段と、

上記入力画像及び上記モデル画像の各々について、上記特徴点の近傍領域にお

ける特徴量を抽出して保持する特徴量保持手段と、

上記入力画像の各特徴点と上記モデル画像の各特徴点とを比較し、類似する特徴量を有する候補対応特徴点ペアを生成する特徴量比較手段と、

上記候補対応特徴点ペアを用いて上記入力画像上の上記モデルの有無を検出し、上記モデルが存在する場合に該モデルの位置及び姿勢を推定するモデル姿勢推定手段とを備え、

上記モデル姿勢推定手段は、ランダムに選択した3組の候補対応特徴点ペアから決定されるアフィン変換パラメータをパラメータ空間に投射する操作を繰り返し、パラメータ空間上で形成されたクラスタのうち最多メンバ数を持つクラスタに属するアフィン変換パラメータに基づいて上記モデルの位置及び姿勢を決定するアフィン変換パラメータを求める

ことを特徴とするロボット装置。

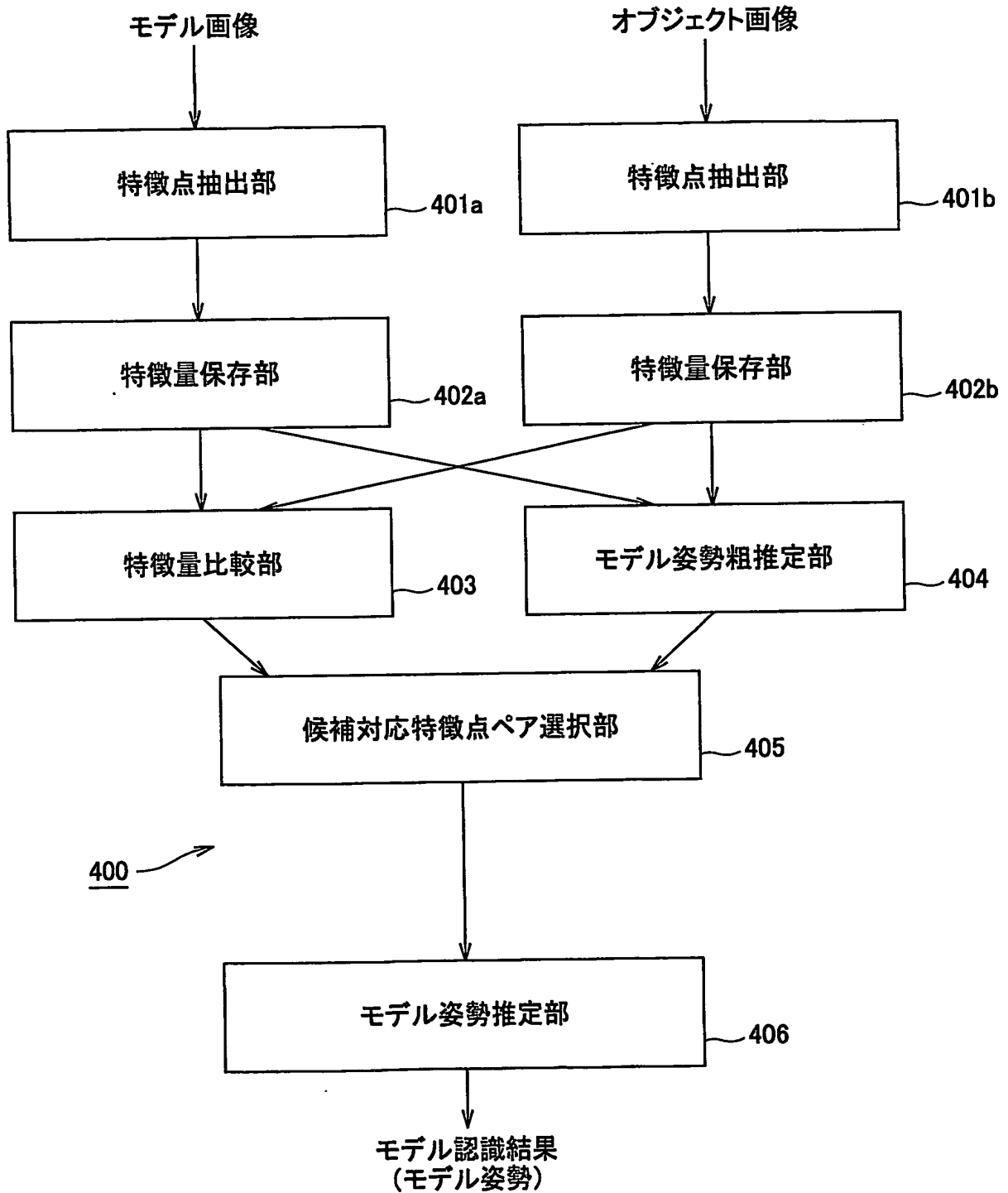


FIG. 1

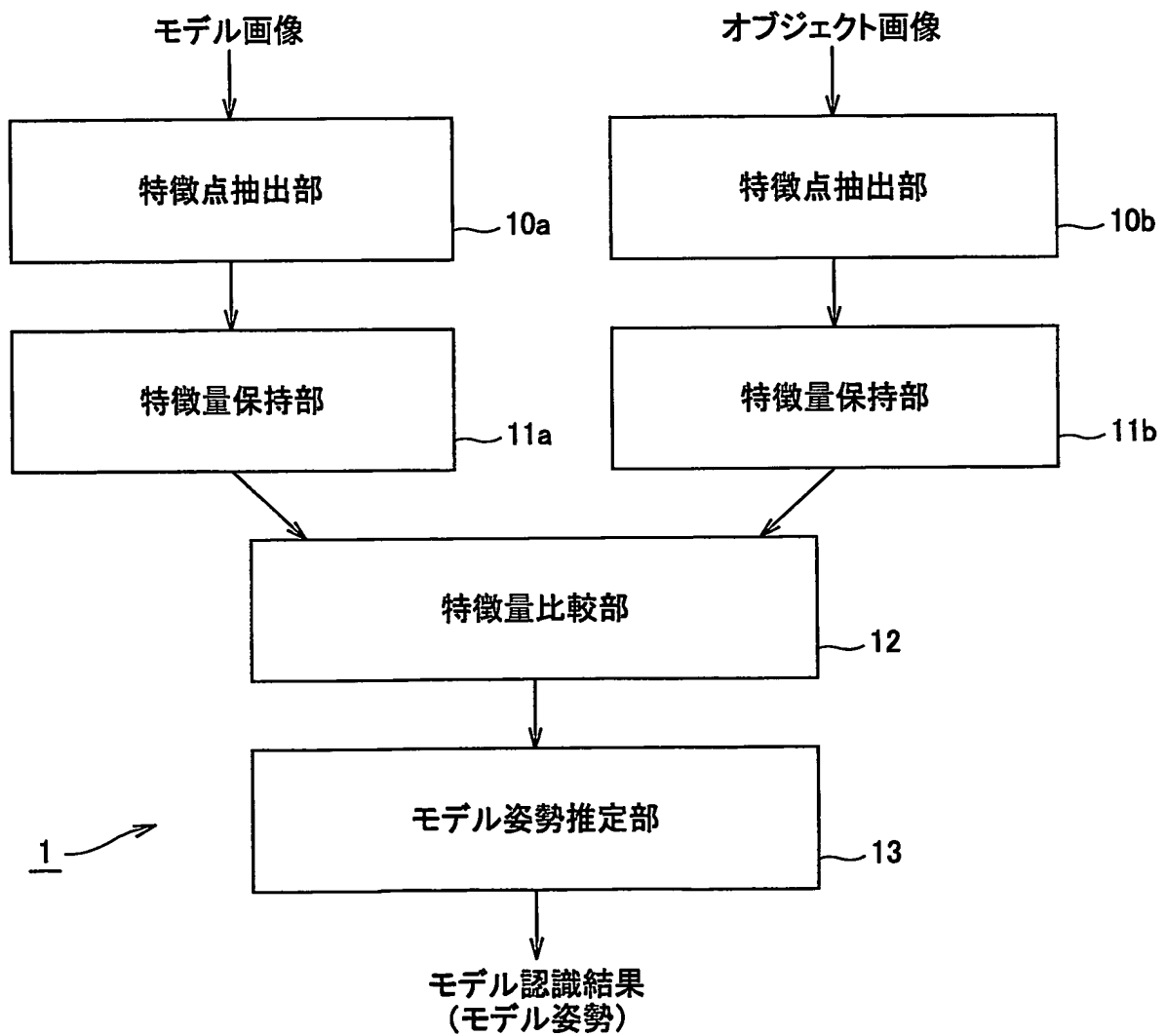


FIG.2

3/14

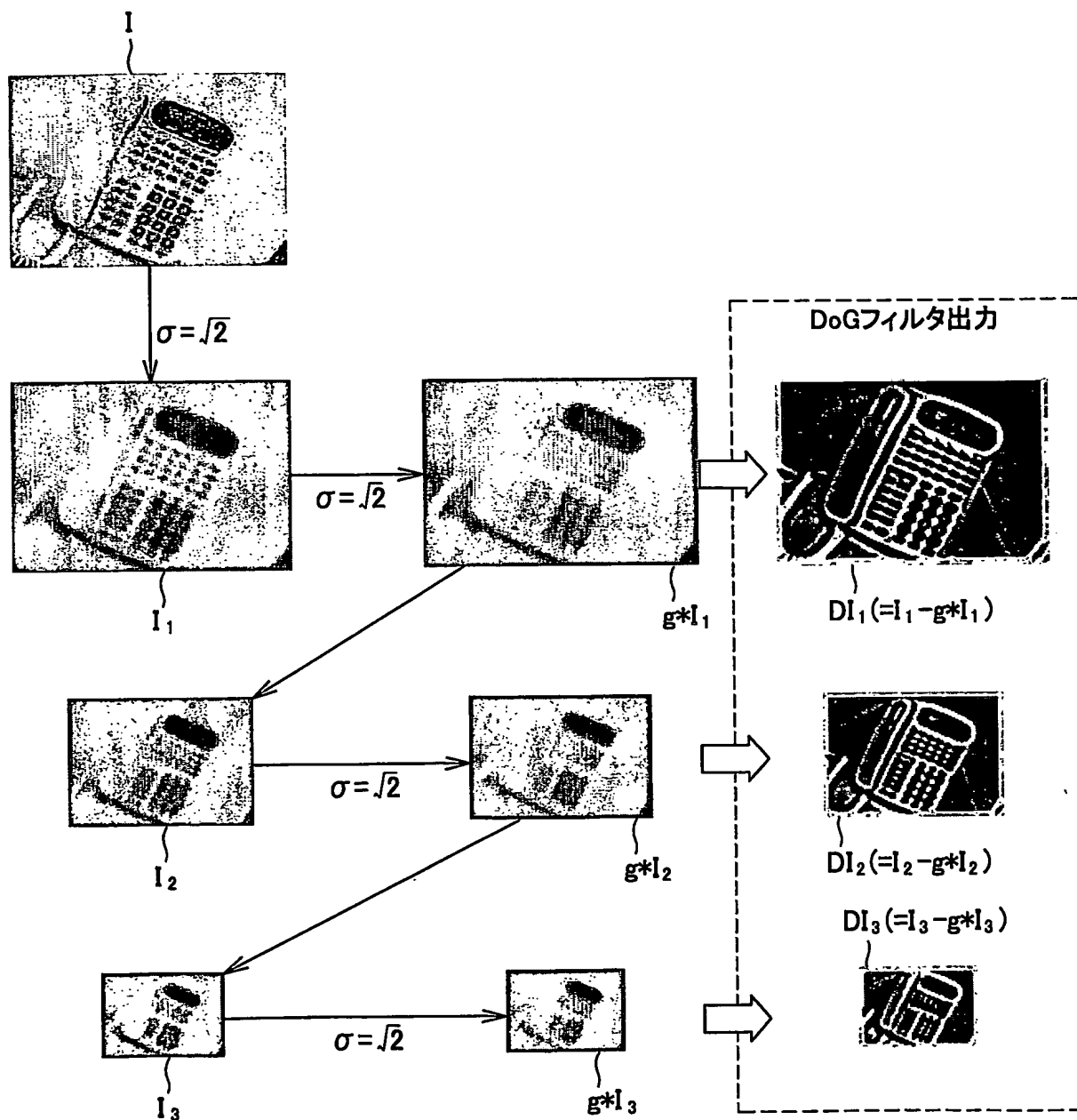


FIG.3



4/14

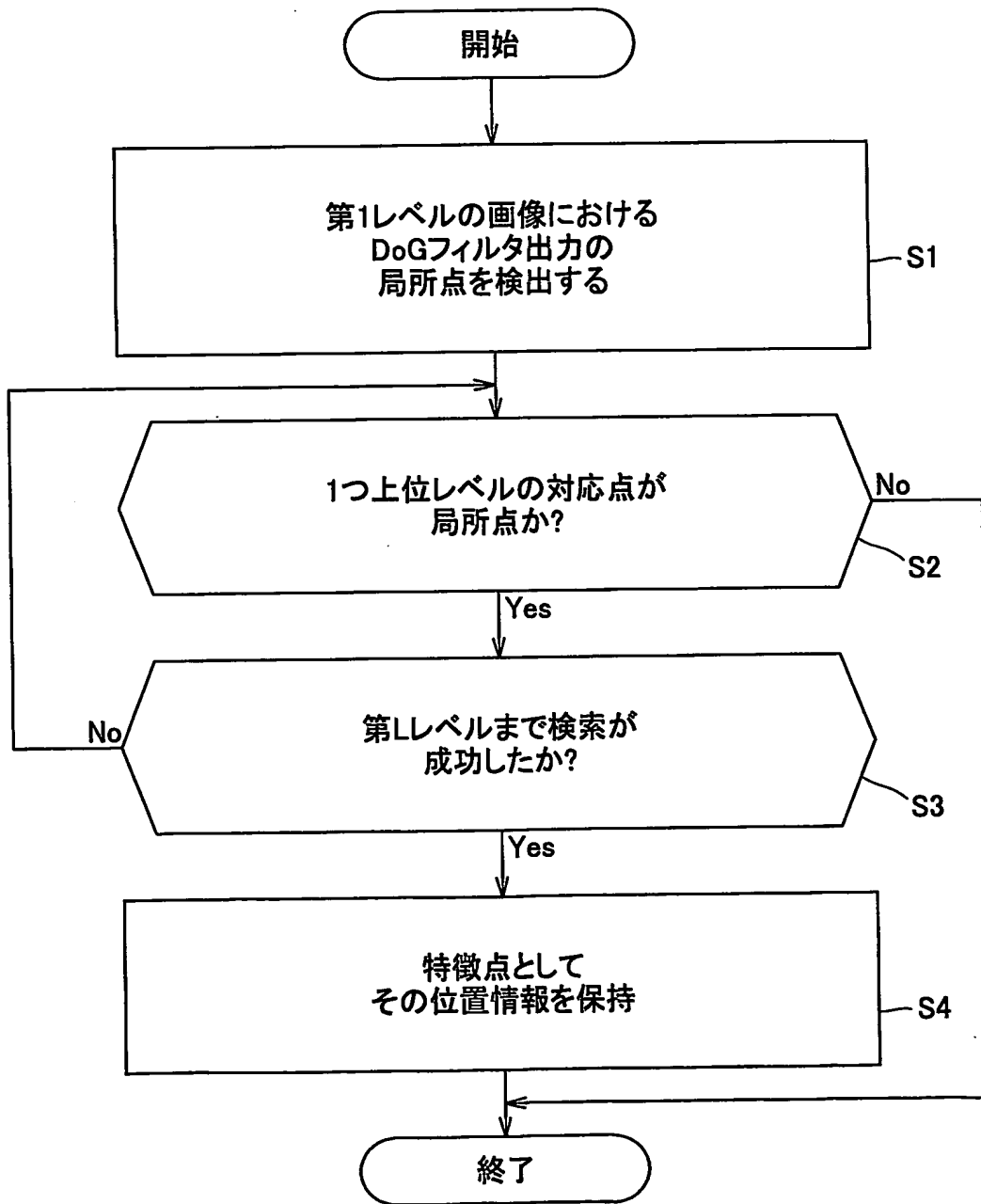


FIG.4

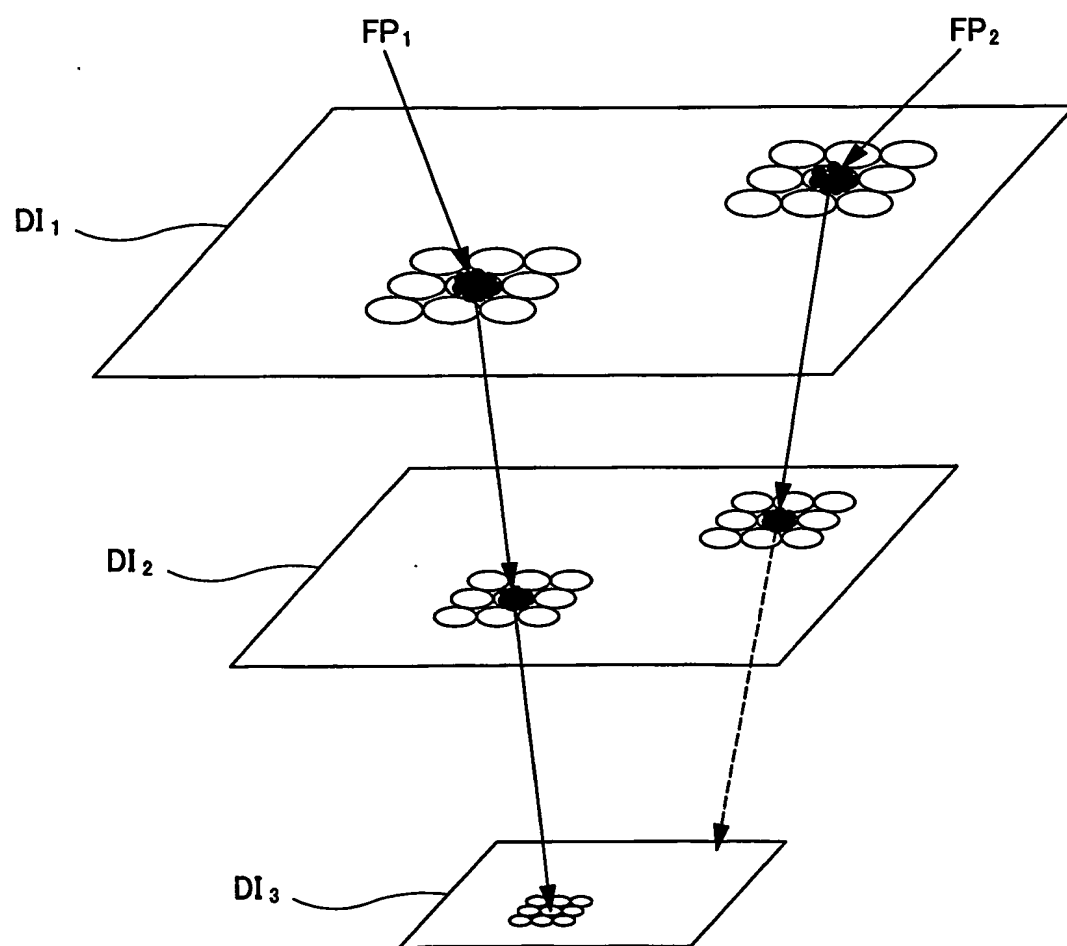


FIG.5

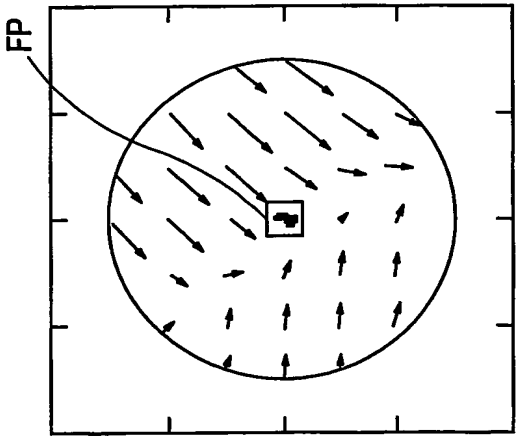


FIG.6A

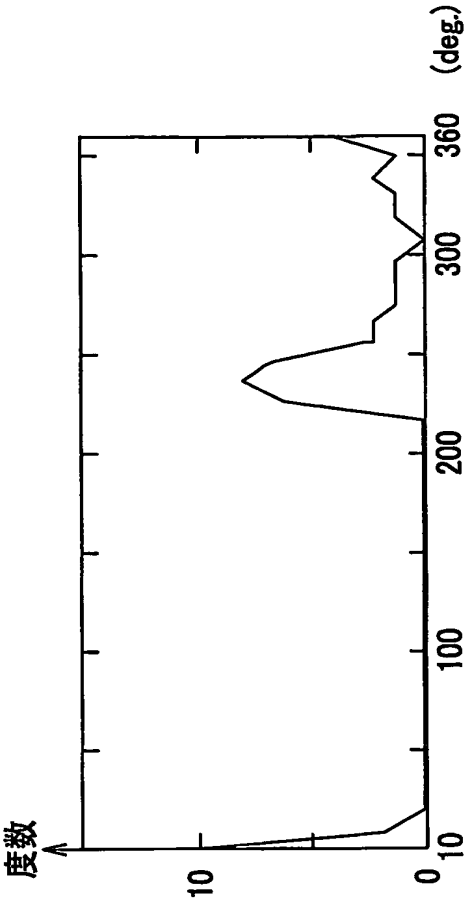


FIG.6B

7/14

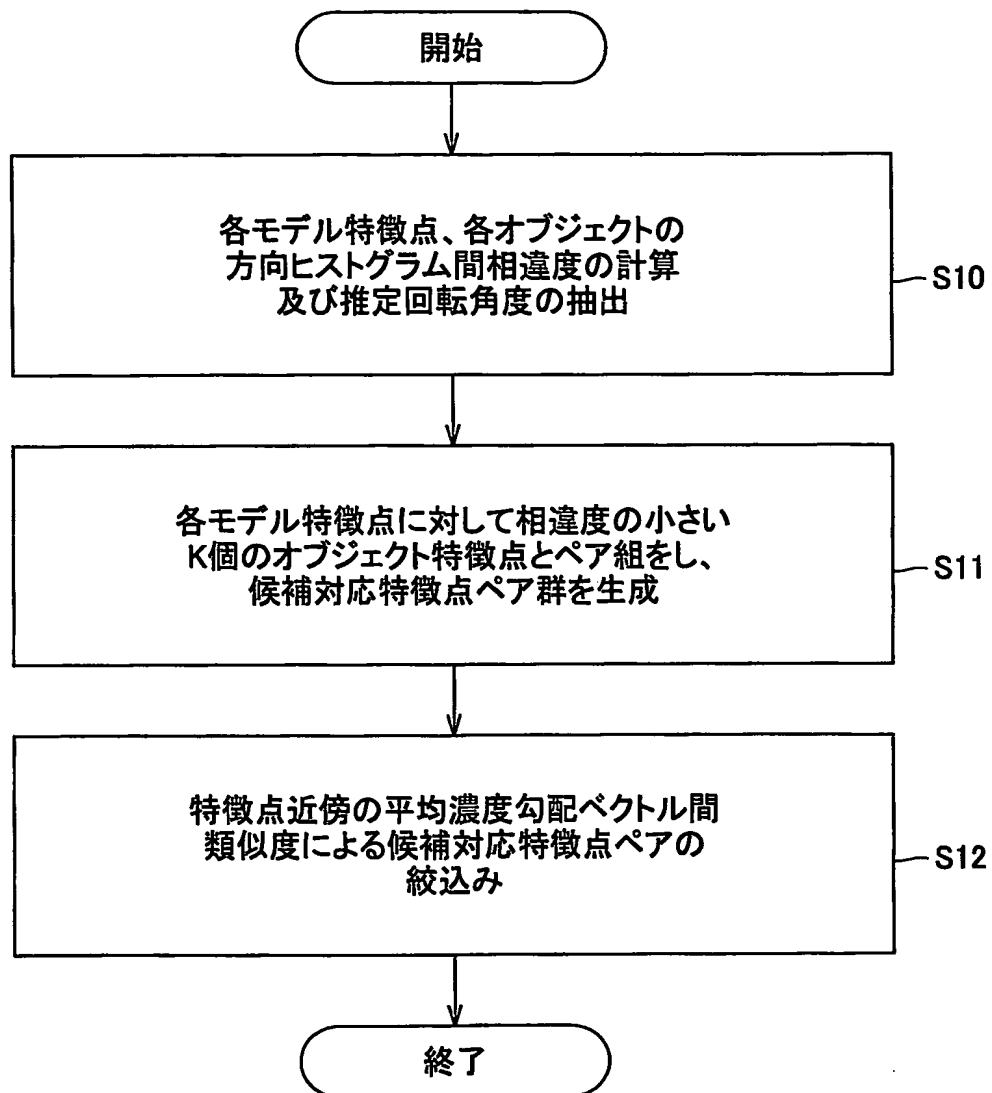


FIG.7

8/14

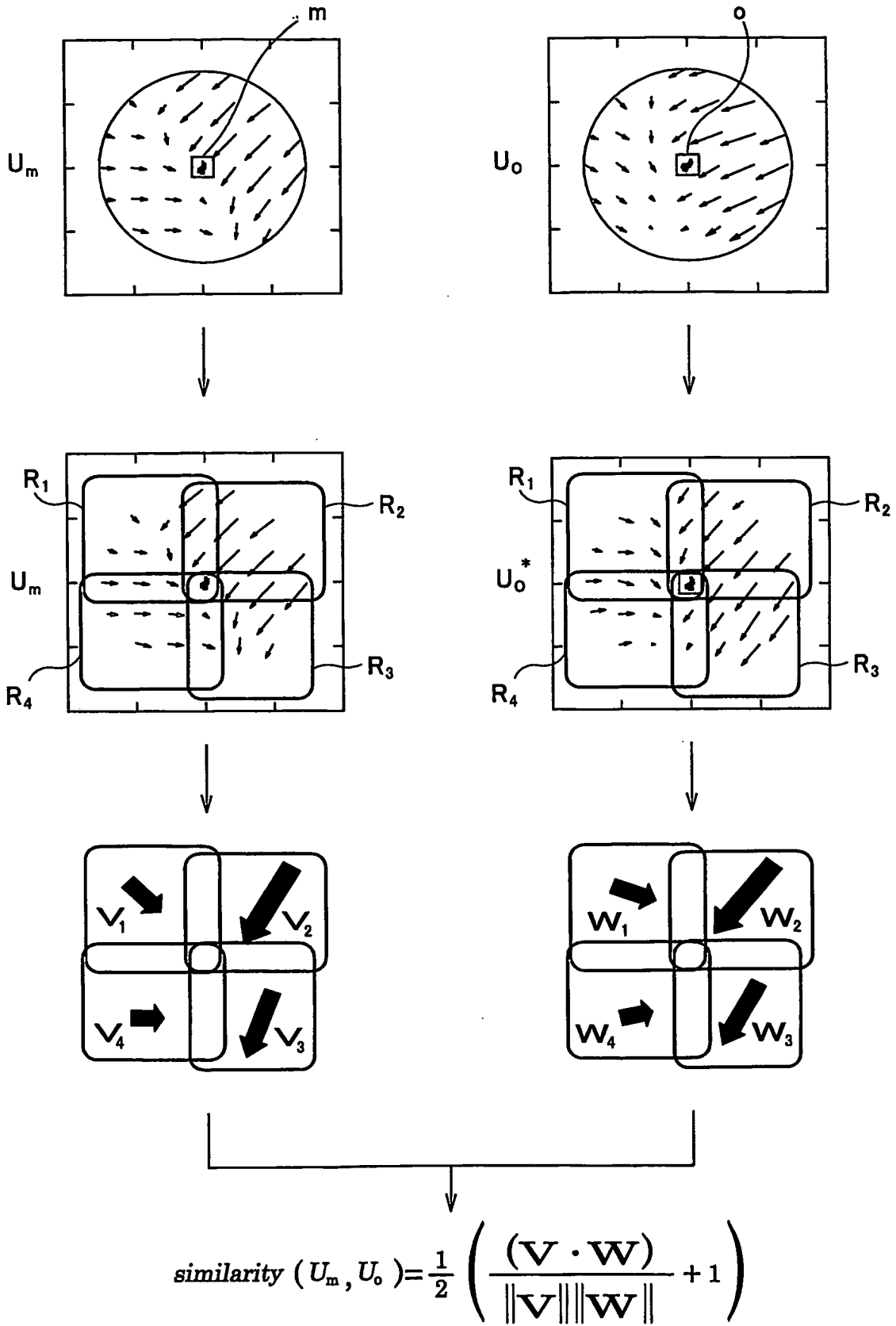


FIG.8

9/14

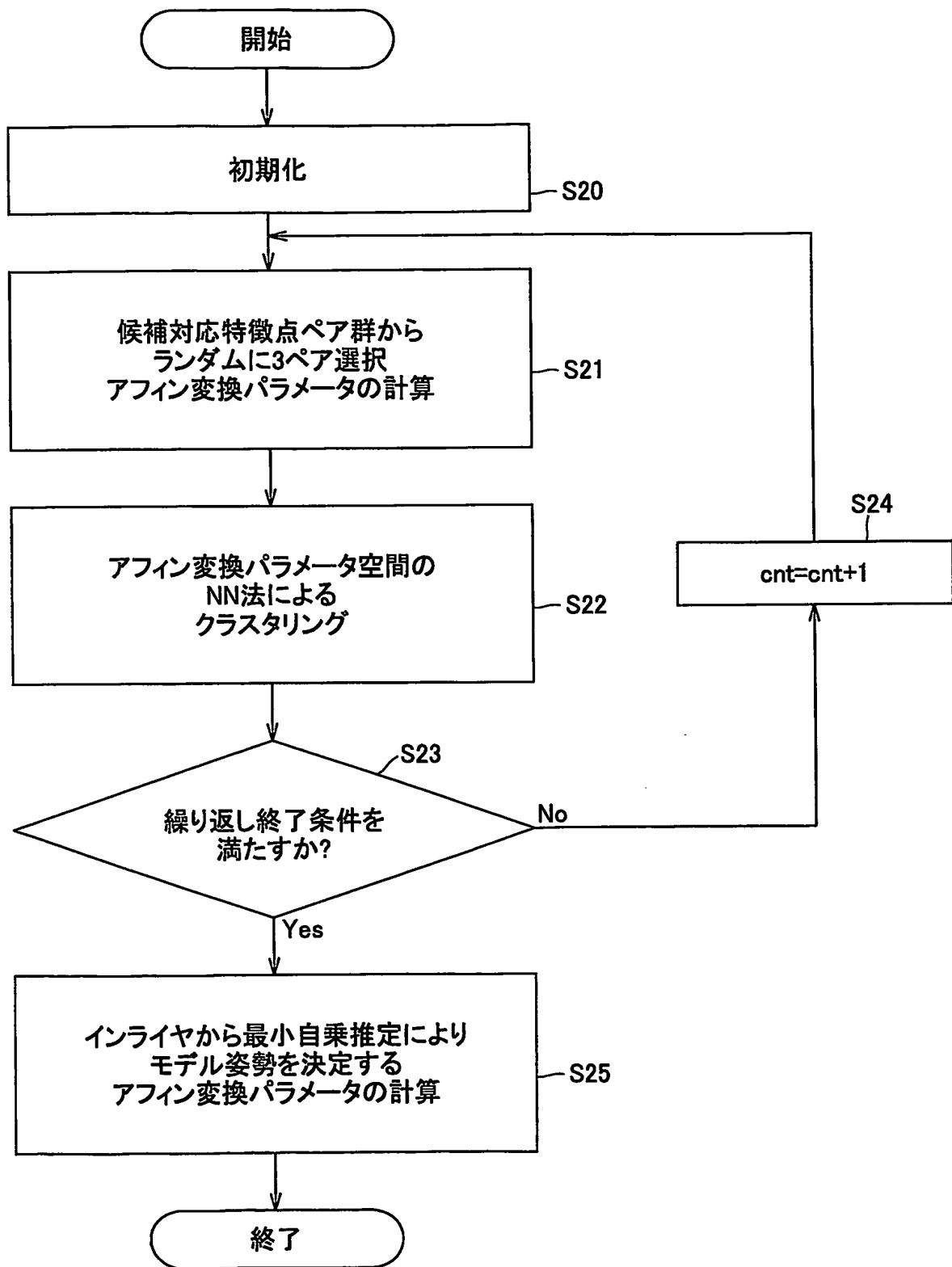


FIG.9

10/14

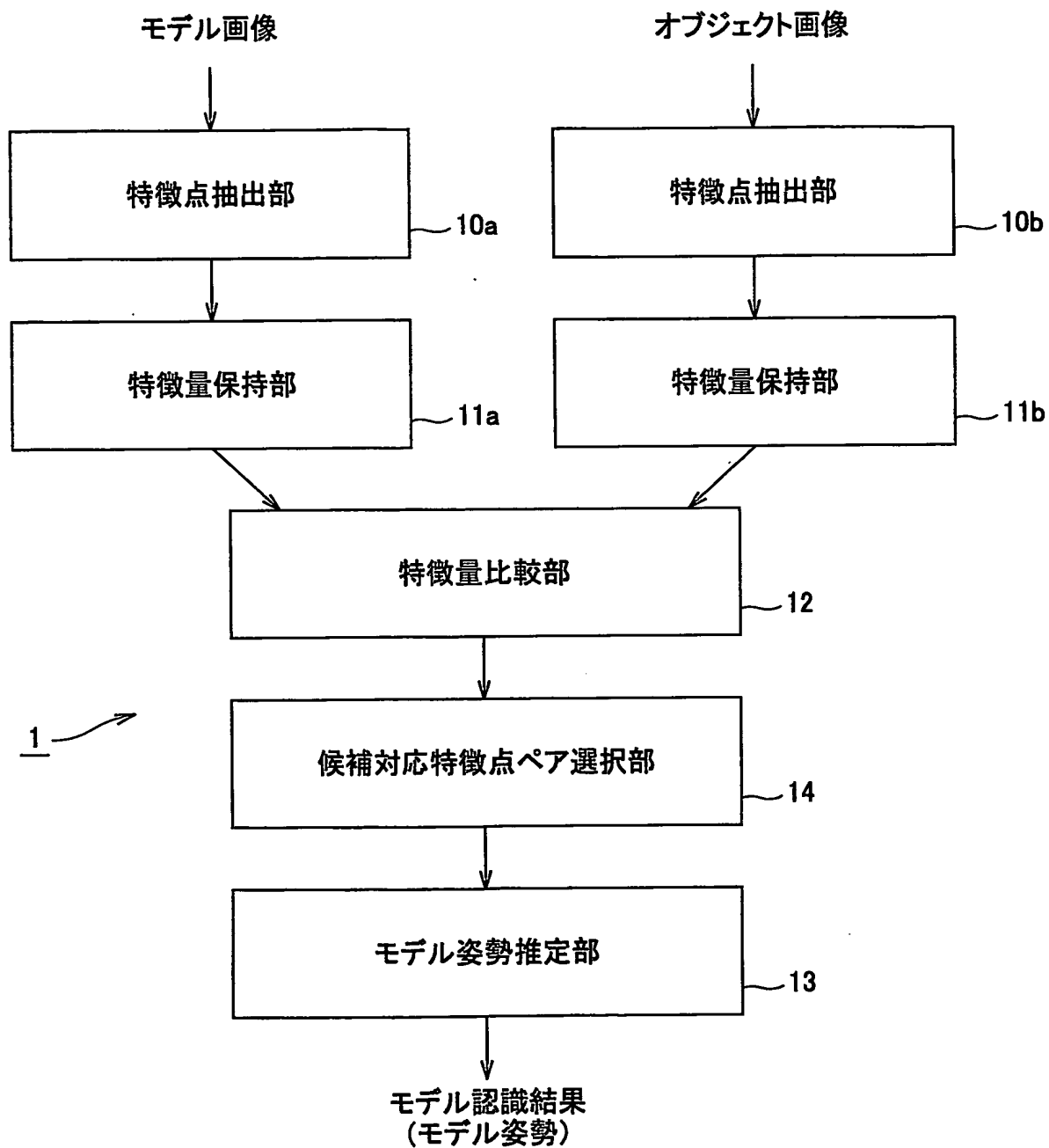


FIG. 10

FIG.1 1A

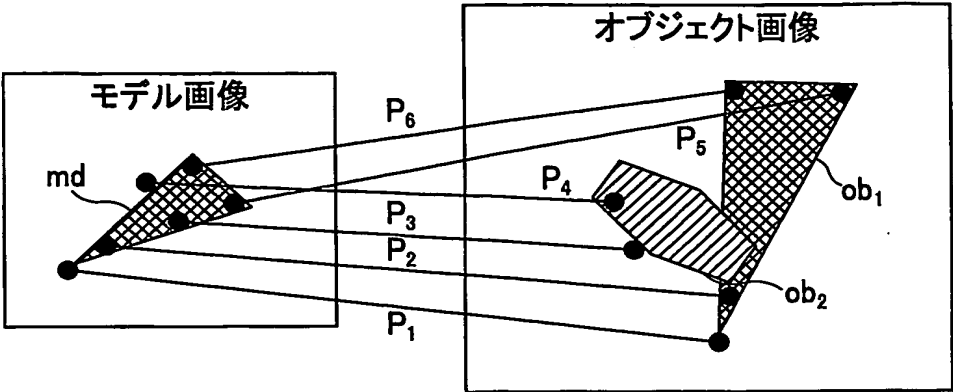


FIG.1 1B

P <sub>1</sub>	(インライヤ)	40deg.
P <sub>2</sub>	(インライヤ)	40deg.
P <sub>3</sub>	(アウトライヤ)	110deg.
P <sub>4</sub>	(アウトライヤ)	260deg.
P <sub>5</sub>	(インライヤ)	40deg.
P <sub>6</sub>	(インライヤ)	40deg.

FIG.1 1C

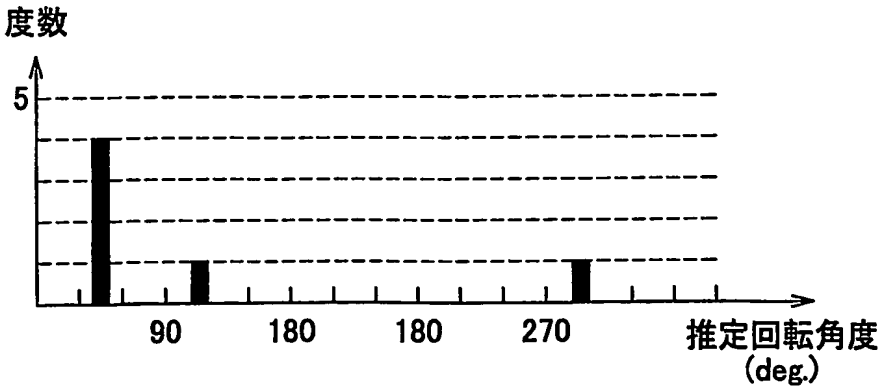
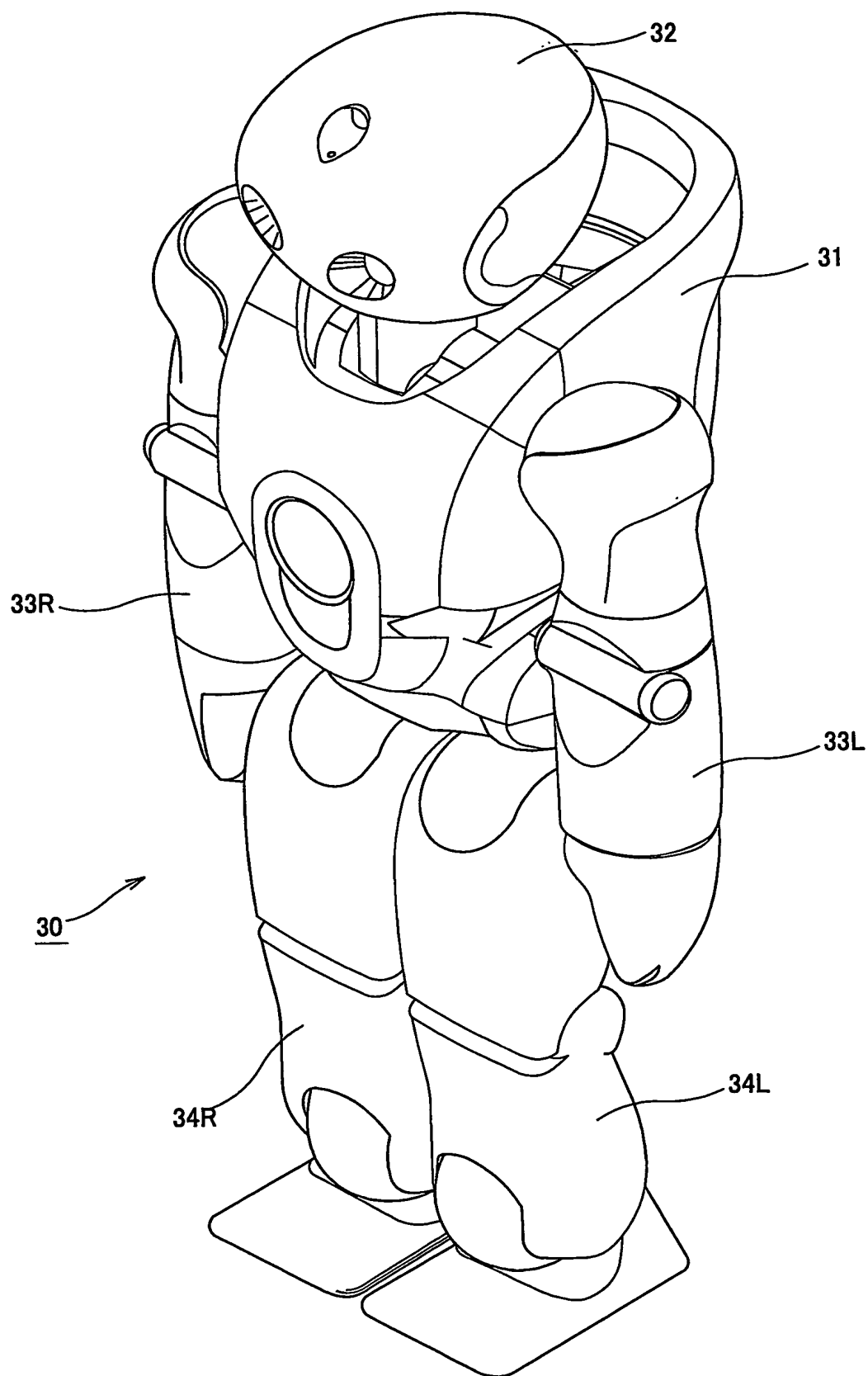


FIG.1 1

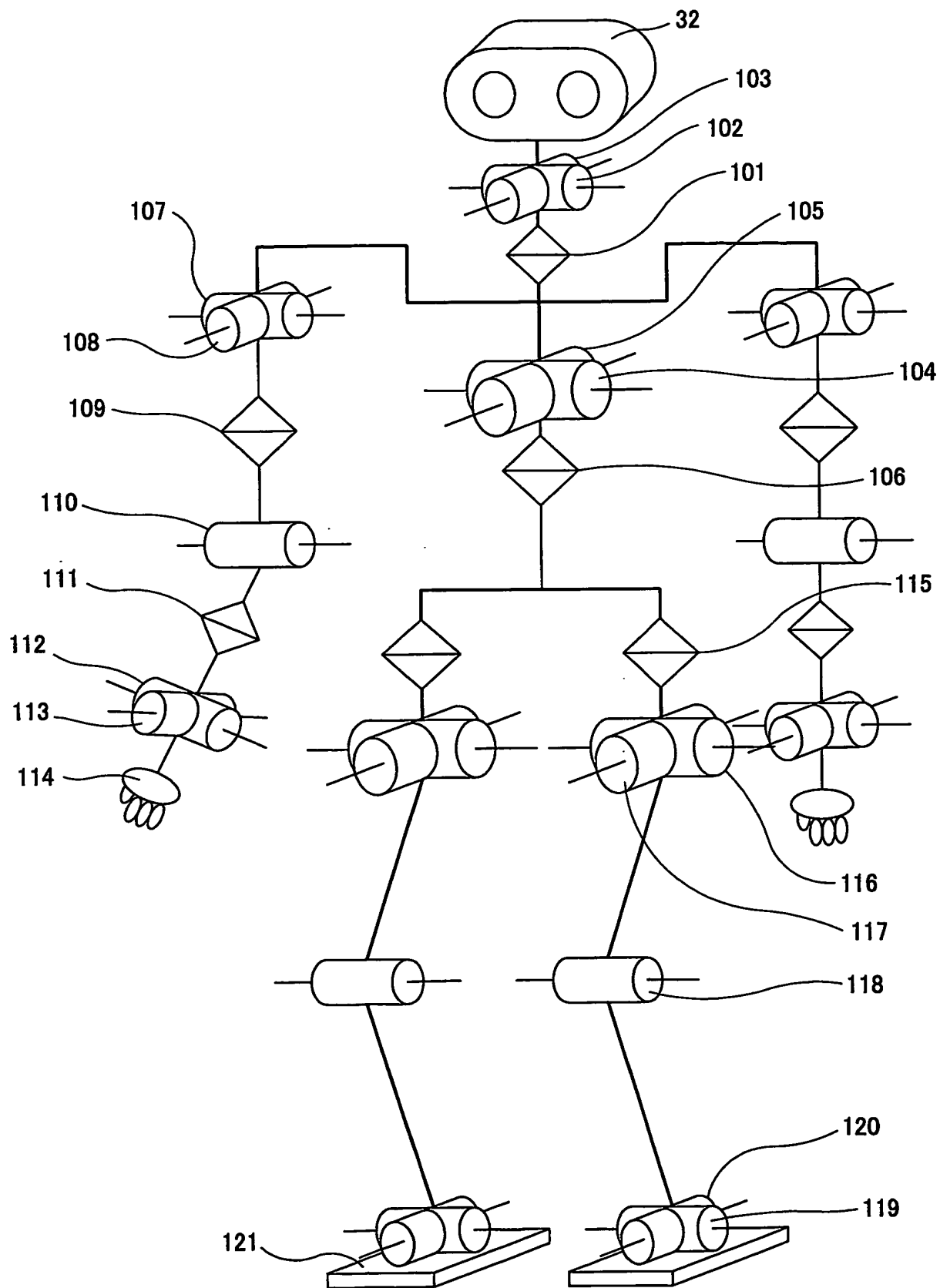


12/14



**FIG.12**

13/14

**FIG.13**

14/14

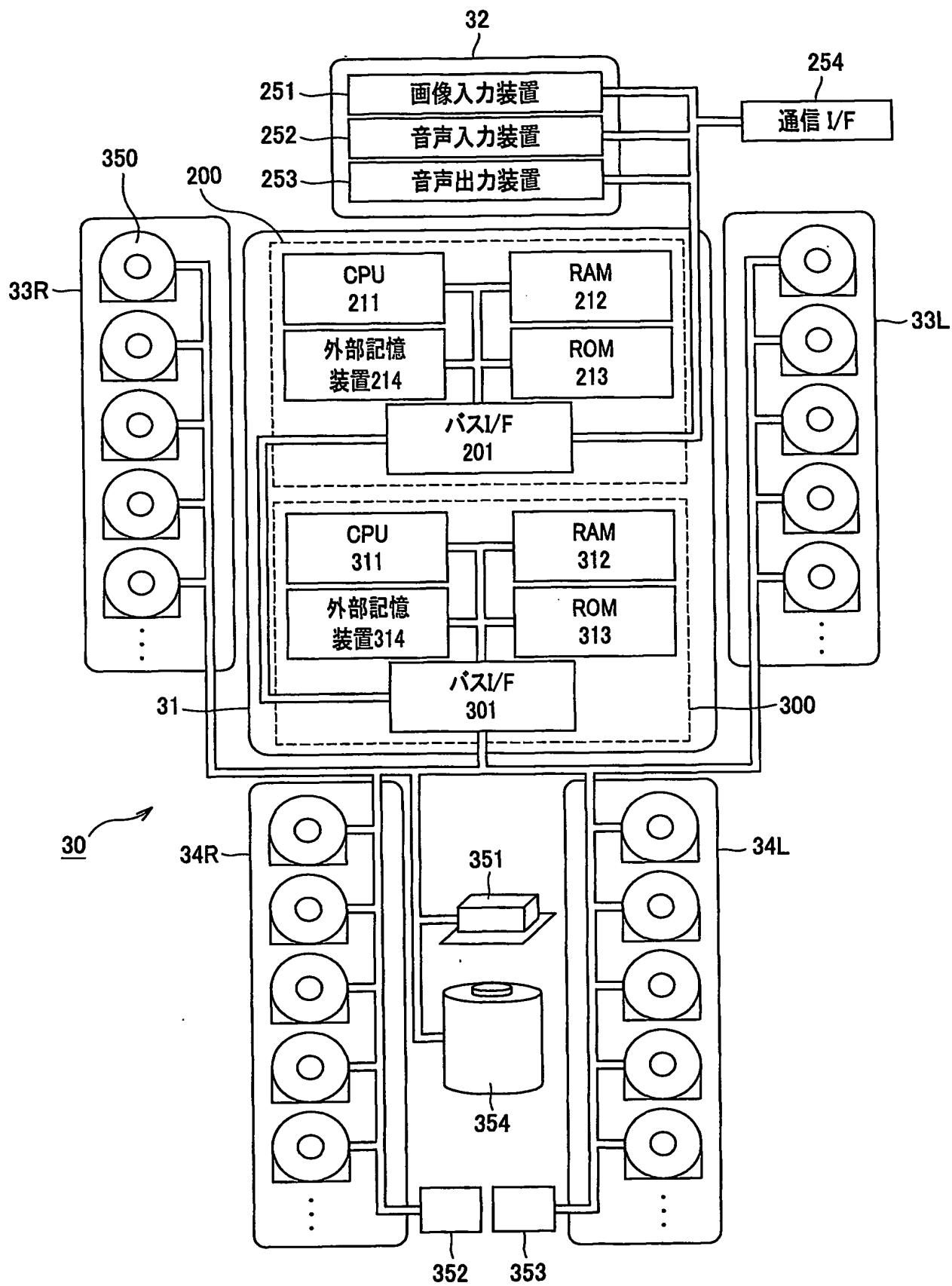


FIG. 14

## INTERNATIONAL SEARCH REPORT

International Application No.

PCT/JP2004/005784

A. CLASSIFICATION OF SUBJECT MATTER  
Int.Cl<sup>7</sup> G06T7/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl<sup>7</sup> G06T7/00-7/60

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2004
Kokai Jitsuyo Shinan Koho	1971-2004	Toroku Jitsuyo Shinan Koho	1994-2004

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	LOWE, D.G, "Object Recognition from Local Scale-Invariant Features", Computer Vision, 1999, The Proceeding of the Seventh IEEE International Conference on, Vol.2, 20-27 September, 1999, pages 1150 to 1157, Vol.2	1, 9, 10, 16, 18
Y	JP 2002-175528 A (Matsushita Electric Industrial Co., Ltd.), 21 June, 2002 (21.06.02), Full text; all drawings (Family: none)	1, 16, 18
Y	JP 2002-8012 A (National Institute of Advanced Industrial Science and Technology), 11 January, 2002 (11.01.02), Full text; all drawings (Family: none)	18

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&amp;" document member of the same patent family

Date of the actual completion of the international search  
26 May, 2004 (26.05.04)Date of mailing of the international search report  
15 June, 2004 (15.06.04)Name and mailing address of the ISA/  
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

## A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl<sup>7</sup> G06T7/00

## B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl<sup>7</sup> G06T7/00 - 7/60

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922-1996年

日本国公開実用新案公報 1971-2004年

日本国実用新案登録公報 1996-2004年

日本国登録実用新案公報 1994-2004年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

## C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	LOWE, D. G, 'Object Recognition from Local Scale-Invariant Features', Computer Vision, 1999. The Proceeding of the Seventh IEEE International Conference on, Volume 2, 20-27 Sept. 1999, pp. 1150-1157 vol.2	1, 9, 10, 16, 18
Y	JP 2002-175528 A (松下電器産業株式会社) 2002. 06. 21, 全文 全図 (ファミリーなし)	1, 16, 18
Y	JP 2002-8012 A (独立行政法人産業技術総合研究所) 2002. 01. 11, 全文 全図 (ファミリーなし)	18

☐ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

## \* 引用文献のカテゴリー

「A」 特に関連のある文献ではなく、一般的技術水準を示すもの

「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)

「O」 口頭による開示、使用、展示等に言及する文献

「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&amp;」 同一パテントファミリー文献

国際調査を完了した日

26. 05. 2004

国際調査報告の発送日

15. 6. 2004

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)

郵便番号 100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

新井 則和

5H

8937

電話番号 03-3581-1101 内線 3531